Microb Ecol (2018) 75:917–929
https://doi.org/10.1007/s00248-017-1087-7

ENVIRONMENTAL MICROBIOLOGY

# Ecological Patterns Among Bacteria and Microbial Eukaryotes Derived from Network Analyses in a Low-Salinity Lake

**Adriane Clark Jones**[1,2] · **K. David Hambright**[3] · **David A. Caron**[1]

Received: 16 December 2016 / Accepted: 9 October 2017 / Published online: 7 November 2017


**Abstract** Microbial communities are comprised of complex assemblages of highly interactive taxa. We employed network analyses to identify and describe microbial interactions and co-occurrence patterns between microbial eukaryotes and bacteria at two locations within a low salinity (0.5–3.5 ppt) lake over an annual cycle. We previously documented that the microbial diversity and community composition within Lake Texoma, southwest USA, were significantly affected by both seasonal forces and a site-specific bloom of the harmful alga, *Prymnesium parvum*. We used network analyses to answer ecological questions involving both the bacterial and microbial eukaryotic datasets and to infer ecological relationships within the microbial communities. Patterns of connectivity at both locations reflected the seasonality of the lake including a large rain disturbance in May, while a comparison of the communities between locations revealed a localized response to the algal bloom. A network built from shared nodes (microbial operational taxonomic units and environmental variables) and correlations identified conserved associations at both locations within the lake. Using network analyses, we were able to detect disturbance events, characterize the ecological extent of a harmful algal bloom, and infer ecological relationships not apparent from diversity statistics alone.

**Keywords** Freshwater ecology · Network analyses · Bacterial community · Microbial eukaryotic community · Protists · Microbial interactions

## Introduction

Microbial communities in aquatic environments perform fundamental ecosystem services, such as primary production, trophic transfer, and nutrient recycling [1, 2]. Interactions between these taxa and higher trophic levels range from mutualistic to predatory or parasitic [3–5]. Central themes in microbial ecology are to quantify and describe patterns of interactions within complex assemblages of microorganisms and to measure and document changes in response to environmental forcing features such as seasonality [6, 7], natural disturbance [8, 9], and experimental manipulation [10, 11].

Network analyses facilitate the exploration of many interconnected correlations simultaneously and have become useful tools for characterizing complex biological systems [12–14]. Various forms of network analysis have documented non-random clusters of microbial taxa, thereby identifying potential metabolic consortia, trophic relationships, keystone species, or modules of co-occurring microorganisms in space and time [15, 16]. These approaches have been applied to examine microbial communities from lakes [17], oceans [18, 19], soil [20–22], and the human microbiome [23]. At the broadest scales, these analyses revealed co-occurrence patterns among microbes within or between ecosystems [16] and helped our understanding of the contribution of specific taxa to global-scale biogeochemical processes, such as the sinking of carbon from surface waters [24].

**Electronic supplementary material** The online version of this article (https://doi.org/10.1007/s00248-017-1087-7) contains supplementary material, which is available to authorized users.

✉ Adriane Clark Jones
ajones@msmu.edu

[1] Department of Biological Sciences, University of Southern California, Los Angeles, CA 90089-0371, USA

[2] Present address: Department of Biological Sciences, Mount Saint Mary's University, Los Angeles, CA 90049, USA

[3] Program in Ecology and Evolutionary Biology, Department of Biology, University of Oklahoma, Norman, OK 73019, USA

In this study, we used network analyses to visualize and quantify individual correlations between microbial metazoa, protists, fungi, bacteria, and environmental parameters in a lake affected by seasonality and ecological disturbance. Lake Texoma is a brackish (PSU 0.5–3.5), temperate lake in the southwest USA that routinely experiences localized harmful algal blooms of the haptophyte *Prymnesium parvum* [25]. *P. parvum* forms ecosystem disruptive algal blooms (EDABs) that are well documented throughout the world [26]. The alga produces a suite of toxins [27, 28], possesses mixotrophic (phagotrophic) ability [29, 30] and thus, can exert strong effects on community structure across many trophic levels [3, 26, 31–35].

We previously characterized the annual cycle of microbial eukaryotic and bacterial diversity at two locations within Lake Texoma: Lebanon Pool and Wilson Creek [36]. In that study, beta diversity (Bray-Curtis similarity matrices) for the bacterial and microbial eukaryotic communities exhibited statistically supported month-to-month seasonal changes and overlaid responses to two specific disturbance events: (1) a biological disturbance in the form of a localized *P. parvum* bloom in Lebanon Pool from January to April, and (2) a physical disturbance in the form of a large spring rain event at both locations during May. The monthly diversity indices exhibited clear patterns of community similarity in response to environmental forcing; however, they did not reveal which specific community members were driving or responding to those changes, nor did they provide insights into the ecological relationships within the microbial food webs [36]. The network approach used here allowed for the simultaneous testing of thousands of taxa-to-taxa correlations and the identification of suites of organisms within the community. The results from Jones et al. 2013 [36] prompted the following research questions centered on disturbance events and taxa-to-taxa ecological relationships: (1) Can network analyses identify and characterize which taxa are positively or negatively associated with ecosystem-disruptive *P. parvum* blooms? (2) Can the network analyses detect seasonal forcing including a onetime physical disturbance event within the year of data? Specifically, which organisms turned over or changed as a result of the disturbance? (3) Do the networks provide a discovery tool to identify consortia of taxa that respond similarly to environmental forcing in Lake Texoma?

## Methods

### Site Description and Sampling

Lake Texoma is a temperate, brackish reservoir on the border of Texas and Oklahoma. Two locations were sampled monthly for 1 year from November 2008 to October 2009: Lebanon Pool and Wilson Creek [36]. Near-shore waters were sampled for temperature, salinity, dissolved oxygen, pH, extracted chlorophyll *a*, *Prymnesium parvum* cell counts, and DNA sequencing of the eukaryotic (18S) and bacterial (16S) rRNA genes. The lake experienced a localized and prolonged bloom of the toxic alga *P. parvum* from January through April in Lebanon Pool only (Fig. S1a-c). Both locations experienced a physical disturbance during May when heavy seasonal rains caused a temporary salinity drop from 3.5 to ≈ 0 ppt. The environmental parameters (temperature, salinity, pH, dissolved oxygen, and nutrients) at the two locations were highly similar and could not alone account for the presence of the *P. parvum* bloom in Lebanon Pool [36].

### Establishing Microbial Operational Taxonomic Units

A detailed description of the protocol for sample processing, sequence generation, and operational taxonomic unit (OTU) calling can be found in Jones et al. [36]. Briefly, the v9 region of the 18S rRNA gene [37] and the v6 region of the 16S rRNA gene [38] were PCR amplified and sequenced via 454 titanium pyrosequencing to assess the diversity of the eukaryotic and bacterial (including plastids) communities. Forty-eight DNA sequence libraries were analyzed (two genes, two locations, 12 months). In the previous study [36], one true replicate (December Wilson Creek) was analyzed and variability between the replicates for each analysis was small compared to the variability between the samples. The open-source software mothur v.1.21.1 (http://mothur.org) was used for sequence processing with the following criteria: (1) an average quality score of 25 or higher, (2) zero ambiguous bases, (3) less than eight homopolymers, (4) exact matches to the primers, and (5) passing chimera check for the 16S rDNA dataset [39]. Mothur v.1.21.1 was used for the alignments using a Needleman-Wunsch algorithm (kmer = 8) with the following criteria: match (+ 1), mismatch (− 1), gap opening (− 2), and gap extension (− 1) to a reference 18S rDNA or 16S rDNA SILVA alignment, and for calling OTUs with the following criteria: 97% similarity using an average neighbor method after a pre-clustering step [39–41]. We found 12,860 OTUs (ranging from 181 to 802 per month) in the 18S dataset and 16,274 OTUs (350 to 771 per month) in the 16S dataset. A taxonomy was assigned to each OTU using a representative sequence (one with the smallest distance to all other sequences within the OTU) and stand-alone BLAST+ [42] to the SILVA small subunit ribosomal database (SSUv108, http://www.arb-silva.de/) [43] and NCBI nucleotide collection nr/nt (www.ncbi.nlm.nih.gov) database. Sequences with high-sequence similarity (at least 80% similarity over 95% of the query) were assigned a taxonomic ID, a term that referred to the classification associated with the sequence that was the best match to our query. Tables S1, S2, and S3 contain a list of OTUs and taxonomic assignments. Table S4 contains a list of OTUs and their representative sequences. The multivariate software

package PRIMER v6.1.7 was used to compute species richness, taxa evenness, and effective diversity of the total microbial communities in each location [44, 45].

### Computing Correlations and Visualizing Connections in Network Diagrams

Our network analyses were built from Spearman's correlations computed over 12 data points (12 months) with each location treated separately. We selected the Spearman correlations (or ranks) for three reasons: (1) to look for monotonic ecological relationships, (2) because a majority of the species were not normally distributed even after transformations were applied, and (3) in an effort to standardize the analyses of the small subunit rDNA genes which can vary greatly in copy number between and within the bacterial and eukaryotic communities. Spearman correlations have also been used in other microbial co-occurrence studies [20, 46]. Only those OTUs detected in at least 4 of the 12 months at a given location and those with relative abundances of at least 0.5% during 1 or more months were included in the analysis. The analysis in Lebanon Pool contained 106 18S OTUs, 128 16S OTUs, and 17 non-OTU or environmental variables. The analysis in Wilson Creek contained 137 18S OTUs, 130 16S OTUs, and 17 non-OTU or environmental variables. We used relative abundances calculated from the total dataset in order to compare the 18S and 16S datasets. We computed alpha diversity indices to test for the likelihood of compositional effects (artefactual correlations) that can be a problem if the sample has a low functional diversity. The relative abundances of the OTUs and absolute quantities of the environmental parameters were normal score transformed as first described in [47]. Pairwise Spearman's correlations (using the average tie-breaking method), with associated permuted $p$ values and $q$ values (false discovery rates) [48], were computed for the time series at each location using the eLSA program and a python script available at http://meta.usc.edu/softs/lsa [49].

Spearman's correlations with $p$ values $\leq 0.01$ (associated $q$ values $< 0.032$ for Lebanon Pool or $< 0.067$ for Wilson Creek) and absolute values of 0.7 were then included in the network analyses. The value of $|0.7|$ was selected to trim the dataset and include only strong correlations. An absolute value of 0.7 is generally considered a meaningful association. The difference in false discovery rates or $q$ values associated with the same $p$ values cutoffs of $\leq 0.01$ reflects the different numbers of pairwise comparisons (31,375 for Lebanon Creek and 40,186 for Wilson Creek) within the two datasets.

Information on variables (nodes) and correlations (edges) were imported into the open-source data viewing platform Cytoscape v.2.8.3 for network analyses and visual representations of all significant correlations (edges) between the bacterial OTUs, eukaryotic OTUs, and variables (nodes) [50–52]. The network nodes (microbial OTUs and environmental variables) were depicted in the network visualizations by shapes and colors, while the network edges (correlations) were depicted as lines in the network diagrams. The total relative abundances for the year of the OTUs were put into eight size bins for visual representations in the network diagrams based on the following criteria: (1) < 0.1, (2) 0.1–0.25, (3) 0.25–0.5, (4) 0.5–0.75, (5) 0.75–1, (6) 1–1.5, (7) 1.5–2, (8) > 2%. The environmental parameters were given the middle size bin of 4.

Three overarching networks were constructed: (a) Lebanon Pool (222 nodes and 4787 correlations), (b) Wilson Creek (247 nodes and 3788 correlations), and (c) a shared dataset, which contained the nodes and correlations (of the same quality [positive or negative], yet not necessarily the same quantity) found at both locations (134 nodes and 510 correlations). The shared dataset was constructed using the advanced network merge Cytoscape plug-in.

Networks were visualized in Cytoscape using two layouts [51]: (1) "edge-weighted, spring-embedded layout" (nodes are strongly repelled or attracted as a function of their Spearman correlation value; spring strength = 50, spring rest length = 100, strength of disconnected spring = 0.05, rest length of disconnected spring = 500, and strength to avoid collisions = 500) and (2) "unweighted, force-directed layout" (the placement of the nodes is optimized based on the number of correlations between the nodes and the value of the correlation is not taken into account; spring coefficient = $1^{-5}$, spring length = 50, node mass = 3).

The networks were undirected, meaning that correlations between nodes did not have directionality. Overarching statistical properties (network density, average clustering coefficient, and average shortest path length) of the networks were computed using the Network Analysis Cytoscape plug-in [53]. The average network density is a normalized value (0 to 1) reflecting the average number of correlations per node and can be used to compare the degree of interconnectivity between different networks of varying sizes. The clustering coefficient (CI) for a node is a ratio (0 to1) relating the number of connections between a node's neighbors (a neighbor is defined as a directly connected node) and the total number of connections possible between all neighbors of that node. Values closer to 1 indicate a highly interconnected network, one in which every node is correlated and connected to every other node. The shortest path length (L) for a node represents the average number of connections required for that node to be connected to all other nodes via the smallest number of connections. Values closer to 1 indicate a highly interconnected network, one in which every node is directly connected to every other node. Each network was statistically compared to a self-randomized version and to an identically sized Erdös-Réyni random model [51, 54] using the Random Networks Cytoscape plug-in (http://web.ecs.syr.edu/~pjmcswee) in order to establish that the connections

detected were not generated at random. The log response ratios of the CI and L (observed values: value from random network) were computed in order to normalize our values and compare the properties of our networks to those generated in other studies. Log distributions of the number of correlations per node (also known as degree distributions) of the networks and Erdös-Réyni random models were plotted and trend lines fitted using SigmaPlot v11.0. The AllegroMCODE Cytoscape plug-in [55] was used to find and isolate highly intercorrelated clusters of nodes within the larger networks. The MCODE parameters were degree cutoff = 2, node score cutoff = 0.03, *K*-score = 2, and max depth 100.

## Results

### Correlation and Network Analyses Revealed Highly Interconnected Microbial Systems

A total of 31,375 (Lebanon Pool) and 40,186 (Wilson Creek) pairwise Spearman correlations were generated (see "Methods" section). The histograms of the associated *p* values (Fig. S2) lacked a uniform distribution, as would be expected under a null model, and instead were skewed towards zero implying meaningful associations [47]. The Spearman correlations with $p$ values $\leq 0.01$ (associated $q$ values $< 0.032$ for Lebanon Pool or $< 0.067$ for Wilson Creek) and absolute values of $> 0.7$ were included in the network analyses. This threshold retained 15% of the correlations (4787 [2669 positive correlations and 2118 negative correlations] between 222 variables) in Lebanon Pool and 9% (3788 [2403 positive correlations and 1385 negative correlations] between 247 variables) in Wilson Creek. The edited dataset for Lebanon Pool contained 94 18S rDNA OTUs including 8 metazoan OTUs, and 122 16S rDNA OTUs including 12 plastid OTUs, while the dataset for Wilson Creek contained 116 18S rDNA OTUs including 6 metazoan OTUs, and 125 16S rDNA OTUs including 11 plastid OTUs. Both datasets included 6 environmental parameters. The shared dataset built from variables and correlations found at both locations contained 134 variables and 510 correlations. Tables S1, S2, and S3 contain a complete list of OTUs, taxonomic assignments, environmental variables, and their attributes.

The number of significant correlations (degrees) per variable (node) ranged from 1 to 111 (Lebanon Pool), 1 to 88 (Wilson Creek), and 0 to 31 (shared network), and did not correlate with an OTU's average relative abundance (Fig. S3), frequency of occurrence (Fig. S4), or taxonomic group (Tables S1 and S2). The microbial community in Wilson Creek had a species richness (Margalef) of 50, a taxa evenness (Pielou) of 0.92, and an effective diversity (Hill) of 151. The edited community in Lebanon Pool had a species richness (Margalef) of 48, a taxa evenness (Pielou) of 0.86, and effective diversity (Hill) of 102. Friedman 2012 recommends an effective diversity of at least 50 to eliminate concern about false correlations resulting from compositional effects [56, 57]. The effective diversity in our samples was high (> 100) indicating a diverse community with meaningful ecological correlations. The distributions representing the frequency of the number of correlations per node for the Lebanon Pool, Wilson Creek, and shared networks (Fig. S4, closed symbols) were each statistically different from the Erdös-Réyni model distributions (Fig. S4, open symbols). The model distributions each fit Poisson curves as expected [58]. The variables' degree distributions for the Lebanon Pool and Wilson Creek networks did not fit power curves as found in some [14], but not all [59] biological networks.

The overwhelming majority (100 and 98%) of the variables and correlations within Lebanon Pool or Wilson Creek each resolved into large highly interconnected networks (Fig. 1 a, e). A majority (70%) of the nodes and correlations shared at both locations also resolved into one interconnected shared network (Figs. S5a).

Network analyses provide a picture of all the direct associations within a biological system. Various topological properties, which include the average clustering coefficient and mean shortest path length (see "Methods" section), describe the patterns of connections found within a network. These features can be statistically compared against those from random networks and can be used as points of comparison with other experimental networks. The average clustering coefficients (CI = 0.601, 0.541, and 0.453) and mean shortest path lengths (L = 2.35, 2.62, and 2.75) of the Lebanon Pool, Wilson Creek, and shared networks, respectively, were each significantly different from the values computed from the Erdös-Réyni random models built from the same number of nodes and correlations, and to self-randomized networks (Table 1).

### The Microbial Networks Revealed Unique Communities Along Broad Seasonal Divides Reflective of a Physical Disturbance Event

Edge-weighted, spring-embedded layouts within Cytoscape were used to organize the networks and position the variables based on their correlations. A majority (> 90%) of the negative correlations within the microbial association networks in Lebanon Pool, Wilson Creek, and the shared network separated the nodes into two distinct subclusters (Figs. 1 and S5). A majority (95%) of the positive correlations were contained within one or the other subclusters at each location. Known seasonality patterns of the variables were overlaid on the networks (Figs.1c–h and S5c, d; yellow highlighted symbols) and indicated that the two subclusters were composed primarily of variables with opposite seasonal abundance patterns divided by the rain disturbance in May: 90% of the variables with a November-to-April seasonality comprised one cluster (Figs. 1c, g and S5c; yellow highlighted symbols) and 90% of

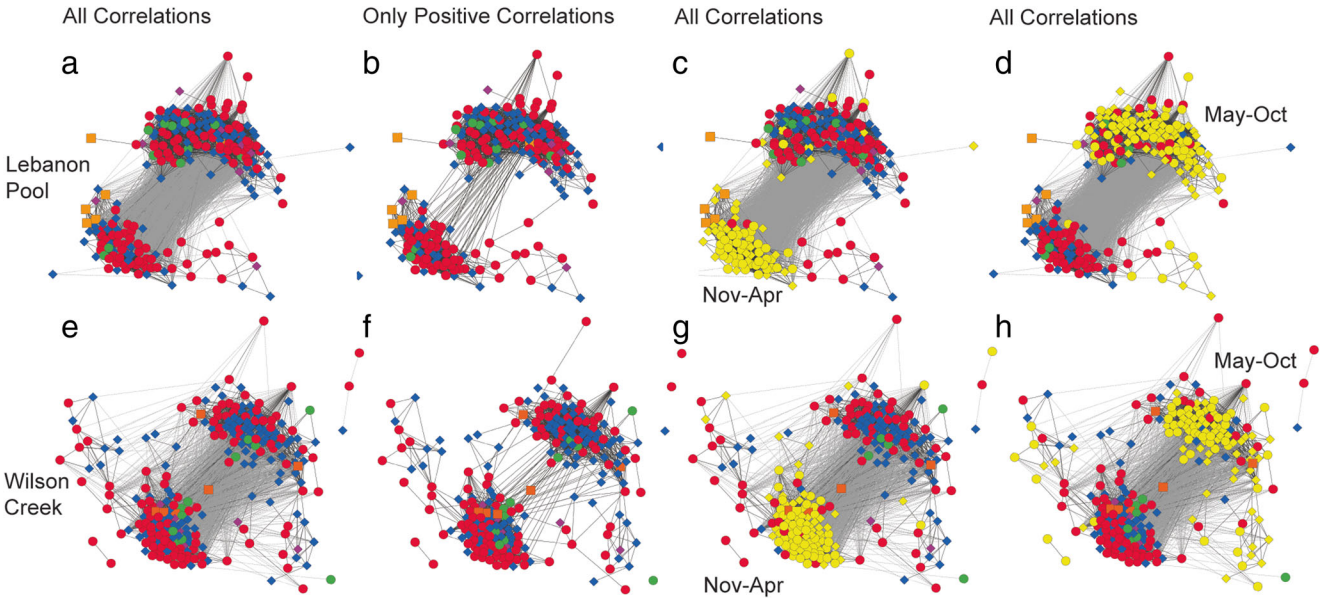


**Fig. 1** The Spearman correlations between the nodes within the network diagrams revealed seasonal abundance patterns among the microbial taxa from Lebanon Pool (**a**–**d**) and Wilson Creek (**e**–**h**). The networks were visualized with the edge-weighted, spring-embedded layout (nodes in the network were positioned based on the strength and sign of their Spearman correlation values). OTUs with 75% or more of their relative abundances contained in the 6-month periods of November–April (**c**, **g**) or May–October (**d**, **h**) are highlighted in yellow. Connections drawn from positive Spearman correlations are black solid lines, and those from negative correlations are gray dotted lines. All correlations (4787 and 3788 [> 0.7 or < − 0.7 and $p$ values ≤ 0.01]) are displayed in **a**, **c**, **d**, **e**, **g**, and **h**. Only positive correlations (2669 and 2403) are displayed in **b** and **f**, demonstrating strong positive correlations between taxa within a season at each location, but not between seasons. Bacteria are red circles, single-celled eukaryotes are blue diamonds, metazoa are purple diamonds, environmental parameters are orange squares, and chloroplasts are green circles

the variables with a May-to-October seasonality comprised the other (Figs. 1d, h and S5d; yellow highlighted symbols).

## Correlations Found Between Algae and Their Plastids Supported the Network Analysis

Positive correlations between plastid OTUs and photosynthetic algal OTUs highlighted likely host/organelle pairs in the networks (Fig. S6). The dataset contained 15 total chloroplast OTUs (11 in Lebanon Pool and 11 Wilson Creek) and all but one (69-U.C.) were identified to phylum level. Thirteen of the chloroplast OTUs showed good correlations to their likely photosynthetic hosts. Two haptophyte plastids (green circles: 12-Hap and 1-Hap) from Lebanon Pool (Fig. S6a) were correlated with each other (0.97), to *P. parvum* cell counts (orange square Ppar) (0.81) and two 18S haptophyte OTUs (blue

**Table 1** Average clustering coefficient (CI) and mean shortest path length (L) for the microbial association networks

| Network type | | Lebanon Pool (222 nodes, 4787 edges) | Wilson Creek (245 nodes, 3788 edges) | Shared (102 nodes, 510 edges) |
|---|---|---|---|---|
| Microbial | CI | 0.601 | 0.541 | 0.453 |
| Erdös-Réyni | CI random | 0.196 (± 0.002) | 0.125 (± 0.001) | 0.10 (± 0.007) |
| | CI/CI random[1] | 3.066 | 4.328 | 4.530 |
| | *ln* r CI[2] | 1.120 | 1.465 | 1.510 |
| Microbial | L | 2.346 | 2.617 | 2.750 |
| Erdös-Réyni | L random | 1.81 (± 0.008) | 1.9 (± 0.009) | 2.24 (± 0.006) |
| | *ln* r L[3] | 0.262 | 0.323 | 0.205 |
| Randomized network | CI random | 0.441 (± 0.010) | 0.318 (± 0.007) | 0.196 (± 0.017) |
| | L random | 1.93 (± 0.004) | 2.1 (± 0.006) | 2.4 (± 0.021) |

[1] CI/CI random: ratio of the CI of the network and the coefficient from an identically sized random network

[2] *ln* r CI: natural log of the ratio

[3] *ln* r L: natural log of the ratio of the L of the network and the coefficient from an identically sized random network values in parentheses is standard deviation

diamonds: 1-Hap and 56-Hap) (0.75 to 0.82). A diatom OTU (blue diamond: 14-Dia) and a diatom plastid (green circle: 10-Dia) had strong correlations in both Lebanon Pool and Wilson Creek, 0.93 and 0.85, respectively (Fig. S6a, b).

### Specific Correlations Between *P. parvum* and Other Taxa in Lebanon Pool

*P. parvum* (18S_OTU1) was the most dominant OTU in Lebanon Pool with a relative abundance of 14% (Table S1) where it formed an ecosystem disruptive harmful algal bloom from January to April (Fig. S1a, b). *P. parvum* (18S_OTU1) had 59 significant correlations (28 negative and 31 positive) in Lebanon Pool (Table S5). Taxa with negative correlations to *P. parvum* included six ciliates (18S_OTU85, 203, 27, 16, 21, and 120), three cryptophytes (18S_OTU31, 47, and 45), and 13 bacterial OTUs. Taxa with positive correlations to *P. parvum* included two fungi (18S_OTU88, and 2), three chrysophytes (18S_OTU114, 108, and 163), three cyanobacteria (16S_OTU39, 62, and 40), and nine alphaproteobacteria (Table S5).

### Positive Correlations Revealed Highly Interconnected and Distinct Microbial Assemblages Within Seasonal Subclusters

The negative correlations between the nodes overwhelmingly separated the microbial communities at each location into two seasonal sub-assemblages (Fig. 1); therefore, we removed the negative correlations from the networks to investigate patterns of positive associations among the taxa. The edge-weighted, spring-embedded layout (Fig. 2) optimized the placement of the nodes based on the number and strength of the positive Spearman correlation values; thus, nodes with the strongest correlation values were placed more closely together. The resulting visual representations of the positive correlations also separated the nodes based on their seasonal abundance patterns in Lebanon Pool (Fig. 2a) and Wilson Creek (Fig. 2b). Clusters on the left were mostly comprised of variables with high abundances during November–April and had network density scores of 0.3 and 0.2 (Lebanon Pool and Wilson Creek, respectively), while clusters on the right were predominantly comprised of variables with high abundances during May–October (network density scores of 0.2). The MCODE algorithm identified the densest, most interconnected regions at each location. The top two scoring MCODE clusters at each location fell into each of the seasonal subclusters (CI of 0.8 and 0.9 in Lebanon Pool; Fig. 2a, yellow highlights, and CI of 0.9 and 0.9 in Wilson Creek; Fig. 2b, yellow highlights). The clusters identified with the MCODE algorithm exhibited qualitatively different compositions of variables that reflected community responses to disturbance (*P. parvum* bloom, January–April in Lebanon Pool only) and seasonality (Fig. 3).

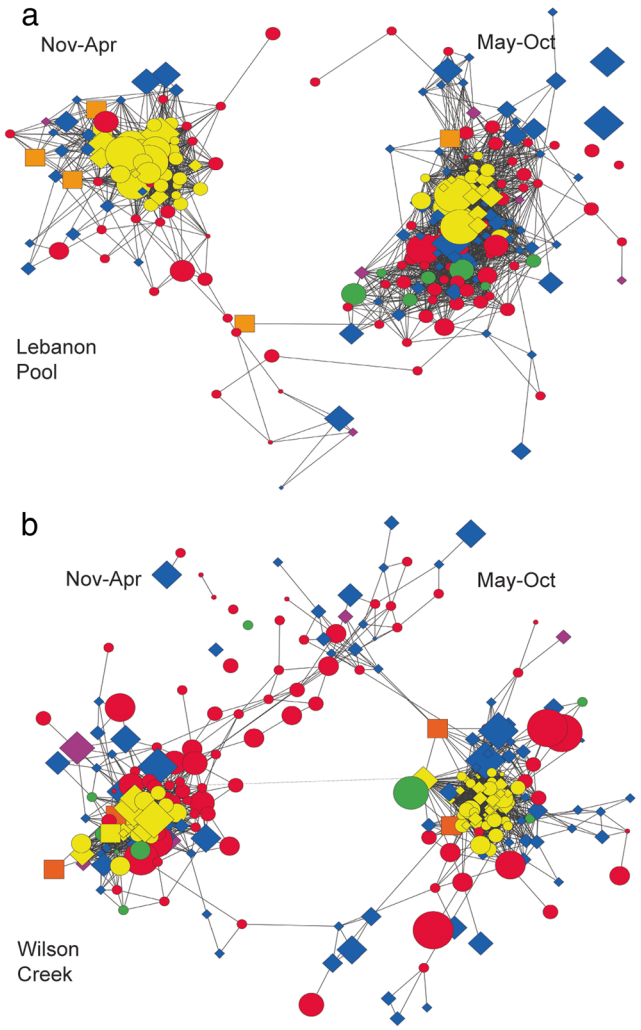


**Fig. 2** Networks representing the positive Spearman correlations between the microbial OTUs and environmental variables from Lebanon Pool (**a**) and Wilson Creek (**b**). The networks were visualized with the edge-weighted, spring-embedded layout (nodes in the network were positioned based on the strength of their positive Spearman correlation values). The visual spread in each network divided the nodes by broad seasonal distribution patterns, the left-hand cluster is primarily composed of OTUs with high abundances from November–April, and the right-hand cluster contains primarily OTUs with higher abundances in May–October. The MCODE algorithm was used to identify regions of highly interconnected clusters of nodes. The top two MCODE clusters in Lebanon Pool (**a**) and Wilson Creek (**b**) are highlighted in yellow. Connections drawn from positive Spearman correlations are black solid lines. Bacteria are red circles, single-celled eukaryotes are blue diamonds, metazoa are purple diamonds, environmental parameters are orange squares, and chloroplasts are green circles. The size of the nodes reflects the average relative sequence abundance

The highly interconnected November–April subcluster in Lebanon Pool contained 41 variables including *P. parvum* counts (orange square Ppar), *Prymnesium* 18S OTUs (blue diamonds: 1-Hap and 56-Hap), and two *Prymnesium*-related plastids (green circles: 1-Hap and 12-Hap); Fig. 3a). These nodes had similar annual relative abundance patterns (Fig.

S1a,c) and were highly interconnected to eight other eukaryotic OTUs (blue diamonds) including two fungal OTUs (2-Fun, 55-Fun), two chrysophyte OTUs (114-Chr and 108-Chr), one chlorophyte OTU (33-Chl), and two unclassified OTUs (32 and 118). The majority (71%) of the cluster was composed of 32 bacterial OTUs (red circles), including four cyanobacteria (a *Synechococcus* [2-Cy] plus filamentous forms [39-Cy, 62-Cy, and 40-Cy]) and 16 (50%) alphaproteobacteria (Alp). Refer to Tables S1 and S2 for a complete list of the OTUs and their identifications.

The highly interconnected November–April cluster in Wilson Creek contained 42 variables and was equally composed of eukaryotic and bacterial OTUs (17 each; Fig. 3c). The eukaryotes included *P. parvum* cell counts (orange square Ppar) and one *Prymnesium* plastid OTU (green circle 1-Hap), although at much lower abundances than for the same period in Lebanon Creek. Similarly, cell counts of *P. parvum* and sequences in *Prymnesium*-related OTUs were much lower in Wilson Creek (Fig. S1a and b compared to c). The eukaryotic community (blue diamonds) in Wilson Creek also included three haptophyte (64-Hap, 56-Hap, 158-Hap), three chlorophyte (Chl), one diatom (134-Dia), four cercozoan (Cer), two ciliate (Cil), and two metazoan OTUs (purple diamonds). The bacterial community (red circles) included a diversity of phylum-level groups, but no cyanobacterial OTUs (Tables S1 and S2).

The highly interconnected communities during May–October (Fig. 3b, d) were comprised of similar numbers of eukaryotic and bacterial OTUs at both locations. The eukaryotic community (blue diamonds) in Lebanon Pool during this period included four cryptophytes (Cry), one chlorophyte (Chl), one diatom, and a diatom plastid (14-Dia, and green circle 10-Dia), three ciliate (Cil), and two metazoan OTUs (purple diamonds: 35-Rot and 281-Rot), while the bacterial community (red circles) was composed of a diversity of taxa, including two cyanobacteria (7-Cy and 51-Cy) (Fig. 3b). The eukaryotic community (blue triangles) in Wilson Creek during this period included three diatom, one cryptophyte, six chlorophyte, three ciliate, and three cercozoan OTUs, while approximately half of the bacterial OTUs in Wilson Creek were cyanobacterial OTUs (red circles Cy; Fig. 3d). The network analyses identified 11 cyanobacterial taxa as a major part of the highly connected cluster in Wilson Creek (Fig. 3d).

### Conserved Correlations in the Shared Network Suggest a Common Microbial Community in the Lake

A total of 134 microbial OTUs, six environmental variables, and 510 interactions (352 positive) were shared between the datasets from Lebanon Pool and Wilson Creek (Table S3 and Figs. 4 and S5a). The negative correlations in the joint network also separated the community into clusters along seasonal divides (Fig. S5c, d). The unweighted force directed layout of the positive correlations also divided the shared community along seasonal divides and into two subclusters (Figs. 4 and S5b). The smaller sub-network observed November–April was 70% composed of bacterial OTUs (red circles) including 10 alphaproteobacteria and four betaproteobacteria, and contained only six shared eukaryotic OTUs (blue diamonds) which included a *Prymnesium* (blue diamond: 1-Hap) and plastid OTU (green circle: 1-Hap) (Fig. 4a). The alphaproteobacteria (red circles Alp) were the most highly connected group and constituted 66% of the correlations during these months. The larger May–October subnetwork contained an equal proportion of eukaryotic (blue diamonds) and bacterial (red circles) OTUs, including four diatoms (Dia), six ciliates (Cil), seven cryptophytes (Cry), and four plastids (green circles) (Fig. 4b). This cluster also contained 11 cyanobacterial OTUs including both *Synechococcus* [Cy(Sy)] and filamentous forms (Cy). The cyanobacteria were the most highly connected group May–October and contained 59% of the total connections.

## Discussion

Network analyses are powerful applications for characterizing complex multi-member biological systems [14]. Microbial assemblages are often comprised of a high diversity of taxa with seemingly overlapping niches that interact in complex ways. Network approaches are a logical tool for quantifying and visualizing microbial occurrence patterns and relationships and, in particular, for analyzing large datasets yielded by high-throughput environmental gene sequencing studies [17, 19–22, 60]. As observed in other microbial co-occurrence networks [6, 7, 19, 46], our analyses identified consortia of taxa that tracked together during the year, implying the existence of physiological/ecological relationships that might explain these co-occurrence patterns. Our networks reflected the seasonality and rain disturbance of the system and organically divided the lake into pre- and post-rain seasonal subcommunities at each location (Fig. 1). The networks also enabled us to more deeply characterize the ecological effects of the *P. parvum* bloom in Lake Texoma (Figs. 2 and 3).

### Strong Associations Implicated by Microbial Networks

Our log response ratios of the average clustering coefficient (CI) (1.1 to 1.5) and the average shortest path lengths (L) (0.21 to 0.32) (Table 1) were comparable to those reported for the multi-domain microbial ecological network from an open-ocean environment [19], presumably indicating similar degrees of connectivity among the microbial taxa in very different planktonic ecosystems. The low average path lengths and high CIs of the networks together suggest that our networks have "small world properties" [61], meaning that the nodes

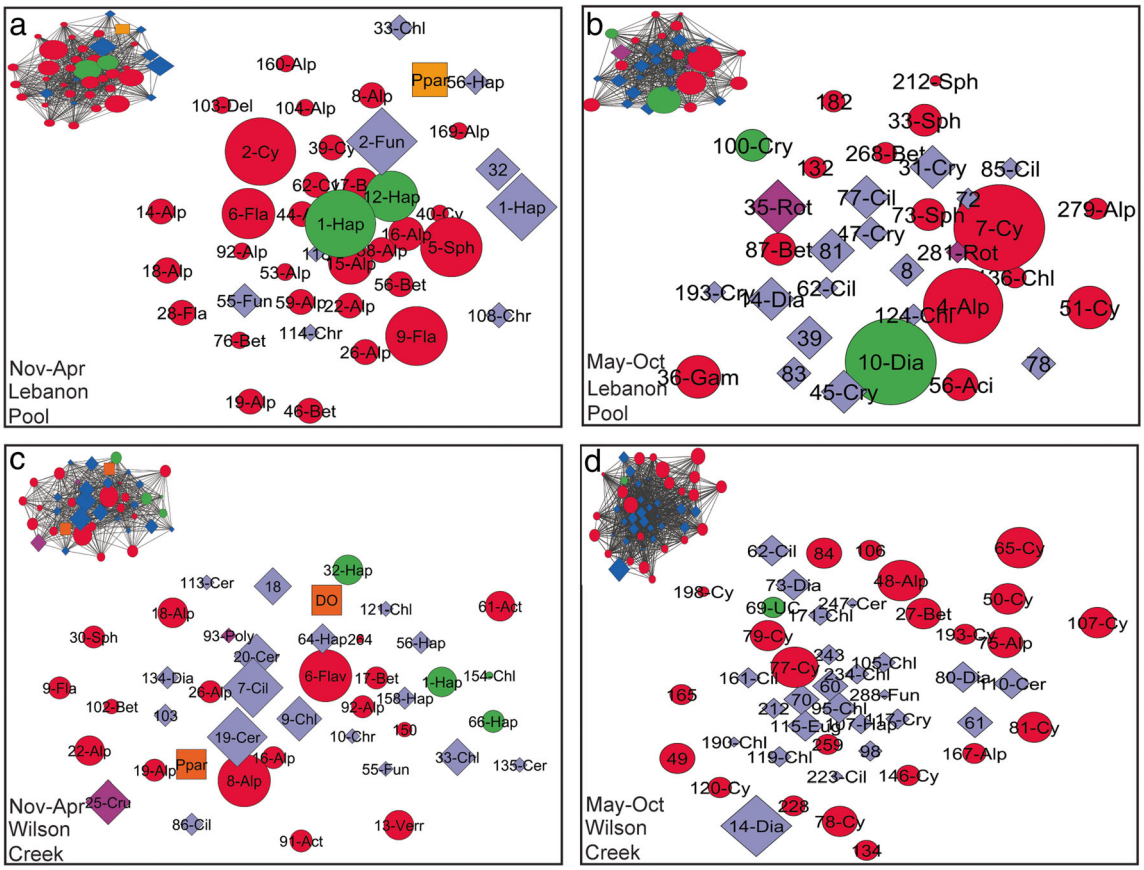


**Fig. 3** Highly interconnected clusters of nodes extracted with the MCODE algorithm from the networks representing all of the positive Spearman correlations from Lebanon Pool (from Fig. 2a) and Wilson Creek (from Fig. 2b). The MCODE algorithm identified two highly interconnected clusters from each location: Lebanon Pool (**a**, **b**) and Wilson Creek (**c**, **d**). Each cluster contained nodes predominantly observed in either November–April (**a**, **c**) or May–October (**b**, **d**). The networks were visualized with the edge-weighted, spring-embedded layout (nodes in the network were positioned based on the strength of their positive Spearman correlation values). The size of the symbols reflects the average relative sequence abundance. Bacteria are red circles, single-celled eukaryotes are blue diamonds, metazoa are purple diamonds, environmental parameters are orange squares [*Prymnesium* cell counts (Ppar) and dissolved oxygen (DO)], and chloroplasts are green circles. Connections drawn from positive Spearman correlations are shown as black solid lines (insets in **a**–**d**). In the larger representations, the connections were omitted for easier reading of the node identifiers. The number on the symbols refers to the OTU identifier numbers. The following identification codes were used for the OTUs with good taxonomic resolution: Alp (alphaproteobacteria), Bet (betaproteobacteria), Gam (gammaproteobacteria), Fla (flavobacteria), Sph (sphingobacteria), Cy (cyanobacteria), Fun (fungi), Hap (haptophyte), Chr (chrysophyte), Chl (chlorophyte), Cer (cercozoa), Cil (ciliate), Cry (cryptophyte), Dia (diatom), Rot (rotifer), Cru (crustacean), Pol (polychaete), and UC (unclassified). Refer to Tables S1 and S2 for a complete list of the OTUs and their identifications

represent highly interconnected consortia of microbes, both eukaryotes and bacteria, with strong associations and interactions.

Our samples were collected at monthly intervals; however, microbes generally interact with each other on much shorter timeframes of hours or days. Thus, our analyses were restricted to contemporaneous associations between consortia of taxa that responded similarly to environmental forcing conditions over the year, and could not reveal time-delayed relationships such as predator-prey or host-parasite. Our datasets contained high functional diversity as described by Friedman and Alm, thus minimizing the risk for artefactual correlations [57]. In addition, strong correlations between plastid OTUs and their eukaryotic hosts provided anticipated and straightforward examples of the ability of the networks to identify biologically meaningful relationships (Fig. S6). Thus, we hypothesize that our networks revealed microbial communities composed of organisms with strong non-random correlations reflecting ecological interactions and coordinated responses to environmental forcing at monthly and seasonal time scales (Table 1 and Supplemental Fig. S4).

## Networks Reflected the Natural History of the Lake

The natural history (strong seasonality, physical rain disturbance, and a toxic algal bloom event) of the two locations in this study provided an interesting framework for our network analyses. Both sampling locations exhibited strong seasonal trends in community composition that reflected forces from gradual seasonal environmental changes, as well as a natural

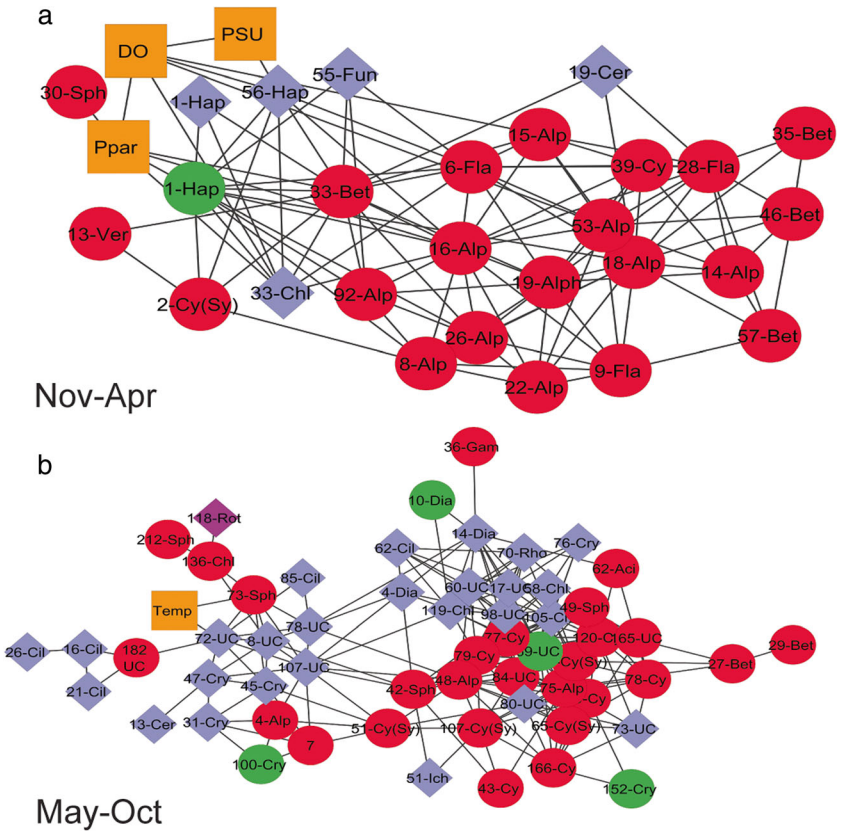


**Fig. 4** Diagrams representing the nodes, and their positive Spearman correlations, shared by Lebanon Pool and Wilson Creek. Network diagrams were visualized with the unweighted force-directed layout (the placement of the nodes was optimized based on the number of correlations between the nodes, and the value of the correlation was not taken into account). The output yielded two discrete clusters: **a** represents the 117 positive correlations across the 38 shared nodes with November–April abundance patterns, while **b** represents the 236 positive correlations across the 64 shared nodes with May–October abundance patterns. Bacteria are red circles, single-celled eukaryotes are blue diamonds, metazoa are purple diamonds, environmental parameters are orange squares [*Prymnesium* cell counts (Ppar), Temperature (Temp), Salinity (PSU), and dissolved oxygen (DO)], and chloroplasts are green circles. Connections drawn from the positive Spearman correlations are connected by black solid lines. The number on the symbols refers to the OTU identifier numbers: Alp (alphaproteobacteria), Bet (betaproteobacteria), Gam (gammaproteobacteria), Fla (flavobacteria), Sph (sphingobacteria), Cy (filamentous cyanobacteria), Cy(Sy) (*Synechococcus* cyanobacteria), Fun (fungi), Hap (haptophyte), Chr (chrysophyte), Chl (chlorophyte), Cer (cercozoa), Cil (ciliate), Cry (cryptophyte), Dia (diatom), Rot (rotifer), Ich (ichthyophonida), and UC (unclassified). The size of the node is not informative. Refer to Table S3 for a complete list of the OTUs and their identifications

physical disruption in the form of a massive spring rain event during May [36].

The seasonality of the system at each location was readily apparent in the force-directed, spring-embedded visualizations of the networks (Figs. 1 and 2). We previously documented seasonal monthly succession of the microbial eukaryotic and bacterial communities likely in response to seasonal fluctuations in temperature, light, and a strong spring rain event [36]. The negative correlations in our networks overwhelmingly reflected opposing seasonal abundance patterns of the OTUs divided by the physical rain disturbance in May (Fig. 1). The networks built only from positive correlations also reflected this seasonal divide as two distinct winter/spring (November–April) and summer/fall (May–October) subnetworks emerged from all of the positive correlations in the lake over the year (Fig. 2). Together, they demonstrate the power of the networks to detect the disturbance event and reveal the qualitative restructuring of the communities at both locations.

### The Ecosystem Disruptive Nature of *P. parvum*

*P. parvum* is classified as an ecosystem disruptive algal bloom (EDAB) [62] and has demonstrated direct negative effects on many members of the plankton community [32, 33, 35]. The alga releases a suite of toxins [27, 63] that can affect several trophic levels ranging from the death of gill-breathing organisms [26], altered life histories and fecundity of crustaceous zooplankton [3, 34], and decreased mobility and/or growth rates of ciliates, flagellated algal cells, diatoms, and filamentous cyanobacteria [31]. In addition, *P. parvum* is mixotrophic, combining phagotrophic and phototrophic nutrition. The alga appears to be an obligate phototroph (i.e.,

requires light for growth), but it also readily engulfs and/or kills a variety of prey ranging from bacteria, to cryptophytes, small diatoms, ciliates, and heterotrophic dinoflagellates [29, 32, 33, 35]. We found that *P. parvum* was negatively correlated with several ciliates and cryptophytes (Table S5), a finding that agrees with published laboratory and field experiments. Interestingly, *P. parvum* was also positively associated with many taxa and one of the highest MCODE scoring subnetworks in Lebanon Pool included *P. parvum*-related nodes (blue diamond: 1-Hap), two plastids (green circles: 1-Hap, 12-Hap), and *P. parvum* cell counts (orange square: Ppar) (Fig. 3a). The highly interconnected consortium in Lebanon Pool yielded a snapshot of the microbial community associated with the harmful alga, which contained only eight other eukaryotic OTUs including two fungi and three chrysophytes. A majority of the bacterial OTUs associated with the *P. parvum* bloom were alphaproteobacteria (Fig. 3a and Table S5). This association in time and space suggests that these taxa may be resistant to the wide-ranging offenses of *P. parvum*. This information provides fodder for hypotheses and the design of future experiments regarding these species' tolerance to *P. parvum.*

The magnitude of the *P. parvum* bloom structured the community in Lebanon Pool and appeared to disrupt the overall diversity of taxa and connections, when compared to the composition of the non-bloom community in Wilson Creek during the same period of time (Fig. 3a compared to c, and [36]). *P. parvum* cells and OTUs were detected during the winter in Wilson creek (Fig. S1b,c), but abundances remained well below bloom conditions, apparently confirming the importance of environmental factors as an important aspect for the success of this species [64, 65]. The most highly connected microbial community within Wilson Creek during the winter/spring contained 17 eukaryotes including cercozoa, ciliates, chlorophytes, and a crustacean metazoan, as well as a diversity of bacterial taxa (Fig. 3c). The strong seasonality of the system makes the direct ecological interpretation of negative correlations with *P. parvum* challenging; however, *P. parvum* was negatively correlated with taxa detected in Wilson Creek including six ciliates and three cryptophytes (Table 5S). Our findings during a natural bloom reinforce the shifts in microbial community structure found by Acosta et al. [66] during an experimental *P. parvum* bloom. Their mesocosm experiments of community structure during bloom and non-bloom conditions revealed that chrysophytes and fungi were strongly associated with the *P. parvum* bloom conditions, while the non-bloom treatments were characterized by diatoms, ciliate, and chlorophytes [66]. Our results also agree with Michaloudi et al. who examined the response of the microbial community during a natural *P. parvum* bloom and found a marked decrease in diatoms and cryptophytes [67].

The disruptive/toxic nature of *P. parvum* was apparent within the properties of the network obtained for Lebanon Pool. The biological disruption that occurred at that site was presumably the reason for the smaller winter/spring subnetwork (75 nodes), compared to the summer/fall (145 nodes) subnetwork (Fig. 2a, left grouping compared to right). In contrast, the seasonal subnetworks in Wilson Creek were similar in size (121 and 125 nodes; Fig. 2b). The log response ratio of the clustering coefficient in Lebanon Pool (Table 1) was lower (1.12) than the ratio in Wilson Creek (1.47) or the shared network (1.51), indicating that fewer of the neighboring nodes were connected to each other for the entire network. On a local scale, *P. parvum* was part of a highly interconnected assemblage within the Lebanon Pool network (Fig. 3a). It would appear that the presence of *P. parvum* may have resulted in a small number of unique taxa that, as a suite, were not strongly interconnected to the taxa that occurred under non-bloom conditions in Lebanon Pool.

## The Shared Network Revealed a Set of Shared Taxa and Connections in the Lake

Nodes representing identical OTUs or environmental parameters and correlations of the same quality (positive or negative) were represented in the shared network (Figs. 4 and S5). Overall, the alphaproteobacteria formed a highly interconnected consortium common to both locations in the winter/spring (Fig. 4a) with very few eukaryotic connections common to both locations. Several ciliates, cryptophytes, chlorophytes, and cyanobacteria also had shared interactions found in the summer/fall at both locations (Fig. 4b).

The shared relationships suggest a microbial community within the lake that responded in parallel at both sites to seasonal environmental forcing, and which persisted through the radical biological disturbance created by the bloom of *P. parvum* in Lebanon Pool. The shared community was comprised of a small proportion of the total number of taxa detected in this study, and an even smaller proportion of the possible correlations, indicating that highly localized environmental cues (including disturbance events) may control much of the assemblage of dominant taxa and many of their detectable interactions.

Taken together, our network analyses were able to detect disturbance events, characterize the ecosystem disruptive nature of the *P. parvum* bloom, and identify consortia of taxa that responded similarly to environmental cues. We see networks as a powerful tool for asking ecological questions and generating testable hypotheses regarding functional relationships among microbial taxa.

Supplementary information is available at Microbial Ecology's website.

**Acknowledgements** The authors would like to thank James D. Easton, Anne C. Easton, and Richard Zamor for field measurements, sample collection and microscopical counts, and Bruce Roe and Fares Z. Najar for DNA sequencing. All sequences are located in the NCBI short read archive under project BioProject PRJNA195945.

**Funding Information** Funding was provided by a grant from the Oklahoma Department of Wildlife Conservation through the Sport Fish Restoration Program (grant F-61-R) to KDH. Oklahoma Mesonet data are provided courtesy of the Oklahoma Mesonet, which is jointly operated by Oklahoma State University and the University of Oklahoma. Continued funding for maintenance of the observing network is provided by the taxpayers of Oklahoma.

## References

1. Fuhrman JA, Caron DA (2015) Heterotrophic planktonic microbes: virus, bacteria, archaea, and protozoa. In: Yates, MV, Nakatsu, CH, Miller, RV, Pillai, SD (eds.) Manual of environmental microbiology, fourth edition. American Society of Microbiology, pp. 4.2.2–1 - 4.2.2–34
2. Worden AZ, Follows MJ, Giovannoni SJ, Wilken S, Zimmerman AE, Keeling PJ (2015) Rethinking the marine carbon cycle: factoring in the multifarious lifestyles of microbes. Science 347. https://doi.org/10.1126/science.1257594
3. Cole JJ (1982) Interactions between bacteria and algae in aquatic ecosystems. Ann Rev Ecol Syst 13:291–314
4. Sherr EB, Sherr BF (2002) Significance of predation by protists in aquatic microbial food webs. Antonie Van Leeuwenhoek 81:293–308. https://doi.org/10.1023/a:1020591307260
5. Ibelings BW, De Bruin A, Kagami M, Rijkeboer M, Brehm M, Donk EV (2004) Host parasite interactions between freshwater phytoplankton and chytrid fungi (Chytridiomycota). J. Phycol. 40: 437–453. https://doi.org/10.1111/j.1529-8817.2004.03117.x
6. Gilbert JA, Field D, Swift P, Newbold L, Oliver A, Smyth T, Somerfield PJ, Huse S, Joint I (2009) The seasonal structure of microbial communities in the western English Channel. Environ. Microbiol. 11:3132–3139. https://doi.org/10.1111/j.1462-2920.2009.02017.x
7. Gilbert JA, Steele JA, Caporaso JG, Steinbruck L, Reeder J, Temperton B, Huse S, McHardy AC, Knight R, Joint I, Somerfield P, Fuhrman JA, Field D (2012) Defining seasonal marine microbial community dynamics. ISME J 6:298–308. https://doi.org/10.1038/ismej.2011.107
8. Jones SE, Chiu CC, Kratz TK, Wu JT, Shadeand A, Mcmahon KD (2008) Typhoons initiate predictable change in aquatic bacterial communities. Limnol. Oceanogr. 53:1319–1326
9. Vigil P, Countway PD, Rose JM, Gobler CJ, Lonsdale DJ, Caron DA (2009) Rapid shifts in dominant taxa among microbial eukaryotes in estuarine ecosystems. Aq Microb Ecol 54:83–100
10. Kim DY, Countway PD, Gast RJ, Caron DA (2011) Rapid shifts in the structure and composition of a protistan assemblage during bottle incubations affect estimates of total protistan species richness. Microb. Ecol. 62:383–398. https://doi.org/10.1007/s00248-011-9816-9
11. Shade A, Read JS, Youngblut ND, Fierer N, Knight R, Kratz TK, Lottig NR, Roden EE, Stanley EH, Stombaugh J, Whitaker RJ, CH W, McMahon KD (2012) Lake microbial communities are resilient after a whole-ecosystem disturbance. ISME J 6:2153–2167
12. Ings TC, Montoya JM, Bascompte J, Blüthgen N, Brown L, Dormann CF, Edwards F, Figueroa D, Jacob U, Jones JI, Lauridsen RB, Ledger ME, Lewis HM, Olesen JM, Van Veen FJF, Warren PH, Woodward G (2009) Review: ecological networks—beyond food webs. J Animal Ecol 78:253–269. https://doi.org/10.1111/j.1365-2656.2008.01460.x
13. Poulin R (2010) Network analysis shining light on parasite ecology and diversity. Trends Parasitol. 26:492–498. https://doi.org/10.1016/j.pt.2010.05.008
14. Proulx SR, Promislow DE, Phillips PC (2005) Network thinking in ecology and evolution. Trends Ecol. Evol. 20:345–353. https://doi.org/10.1016/j.tree.2005.04.004
15. Fuhrman JA, Cram JA, Needham DM (2015) Marine microbial community dynamics and their ecological interpretation. Nat Rev Micro 13:133–146. https://doi.org/10.1038/nrmicro3417
16. Williams RJ, Howe A, Hofmockel KS (2014) Demonstrating microbial co-occurrence pattern analyses within and between ecosystems. Frontiers Microbiol 5:358. https://doi.org/10.3389/fmicb.2014.00358
17. Eiler A, Heinrich F, Bertilsson S (2012) Coherent dynamics and association networks among lake bacterioplankton taxa. ISME J 6: 330–342. https://doi.org/10.1038/ismej.2011.113
18. Milici M, Deng Z-L, Tomasch J, Decelle J, Wos-Oxley ML, Wang H, Jáuregui R, Plumeier I, Giebel H-A, Badewien TH, Wurst M, Pieper DH, Simon M, Wagner-Döbler I (2016) Co-occurrence analysis of microbial taxa in the Atlantic Ocean reveals high connectivity in the free-living bacterioplankton. Frontiers Microbiol 7:649. https://doi.org/10.3389/fmicb.2016.00649
19. Steele JA, Countway PD, Xia L, Vigil PD, Beman JM, Kim DY, Chow C-ET, Sachdeva R, Jones AC, Schwalbach MS, Rose JM, Hewson I, Patel A, Sun F, Caron DA, Fuhrman JA (2011) Marine bacterial, archaeal and protistan association networks reveal ecological linkages. ISME J 5:1414–1425
20. Barberan A, Bates ST, Casamayor EO, Fierer N (2012) Using network analysis to explore co-occurrence patterns in soil microbial communities. ISME J 6:343–351. https://doi.org/10.1038/ismej.2011.119
21. de Menezes AB, Prendergast-Miller MT, Richardson AE, Toscas P, Farrell M, Macdonald LM, Baker G, Wark T, Thrall PH (2015) Network analysis reveals that bacteria and fungi form modules that correlate independently with soil parameters. Environ. Microbiol. 17:2677–2689. https://doi.org/10.1111/1462-2920.12559
22. Lupatini M, Suleiman A, Jacques R, Antoniolli Z, Ferreira A, Kuramae EE, Roesch L (2014) Network topology reveal high connectance levels and few key microbial genera within soils. Frontiers Environ Sci 2. https://doi.org/10.3389/fenvs.2014.00010
23. Faust K, Sathirapongsasuti JF, Izard J, Segata N, Gevers D, Raes J, Huttenhower C (2012) Microbial co-occurrence relationships in the human microbiome. PLoS Comp Biol 8:e1002606. https://doi.org/10.1371/journal.pcbi.1002606
24. Guidi L, Chaffron S, Bittner L, Eveillard D, Larhlimi A, Roux S, Darzi Y, Audic S, Berline L, Brum J, Coelho LP, Espinoza JCI, Malviya S, Sunagawa S, Dimier C, Kandels-Lewis S, Picheral M, Poulain J, Searson S, Tara Oceans Consortium C, Stemmann L, Not F, Hingamp P, Speich S, Follows M, Karp-Boss L, Boss E, Ogata H, Pesant S, Weissenbach J, Wincker P, Acinas SG, Bork P, de Vargas C, Iudicone D, Sullivan MB, Raes J, Karsenti E, Bowler C, Gorsky G (2016) Plankton networks driving carbon export in the oligotrophic ocean. Nature 532:465–470. https://doi.org/10.1038/nature16942
25. Hambright KD, Zamor RM, Easton JD, Glenn KL, Remmel EJ, Easton AC (2010) Temporal and spatial variability of an invasive toxigenic protist in a North American subtropical reservoir. Harmful Algae 9:568–577. https://doi.org/10.1016/j.hal.2010.04.006
26. Evardsen B, Imai I (2006) The ecology of harmful flagellates within Prymnesiophyceae and Raphidophyceae. In: Granéli E, Turner J (eds) Ecology of harmful algae. Springer-Verlag, Berlin, pp. 67–79
27. Henrikson JC, Gharfeh MS, Easton AC, Easton JD, Glenn KL, Shadfan M, Mooberry SL, Hambright KD, Cichewicz RH (2010) Reassessing the ichthyotoxin profile of cultured *Prymnesium parvum* (golden algae) and comparing it to samples collected from recent freshwater bloom and fish kill events in North America. Toxicon 55:1396–1404

28. Igarashi T, Satake M, Yasumoto T (1999) Structures and partial stereochemical assignments for prymnesin-1 and prymnesin-2: potent hemolytic and ichthyotoxic glycosides isolated from the red ride alga *Prymnesium parvum*. J. Am. Chem. Soc. 121:8499–8511. https://doi.org/10.1021/ja991740e
29. Tillman U (1998) Phagotrophy by a plastidic haptophyte, *Prymnesium patelliferum*. Aq Microb Ecol 14:155–160
30. Remmel EJ, Hambright KD (2012) Toxin-assisted micropredation: experimental evidence shows that contact micropredation rather than exotoxicity is the role of *Prymnesium* toxins. Ecol. Lett. 15:126–132. https://doi.org/10.1111/j.1461-0248.2011.01718.x
31. Fistarol GO, Legrand C, Graneli E (2003) Allelopathic effect of *Prymnesium parvum* on a natural plankton community. Mar. Ecol. Prog. Ser. 255:115–125
32. Martin-Cereceda M, Novarino G, Young JR (2003) Grazing by *Prymnesium parvum* on small planktonic diatoms. Aq Microb Ecol 33:191–199. https://doi.org/10.3354/ame033191
33. Skovgaard A, Hansen PJ (2003) Food uptake in the harmful alga *Prymnesium parvum* mediated by excreted toxins. Limnol. Oceanogr. 48:1161–1166
34. Nejstgaard JC, Solberg PT (1996) Repression of copepod feeding and fecundity by the toxic haptophyte *Prymnesium patelliferum*. Sarsia 81:339–344. https://doi.org/10.1080/00364827.1996.10413631
35. Tillmann U (2003) Kill and eat your predator: a winning strategy of the planktonic flagellate *Prymnesium parvum*. Aq Microb Ecol 32:73–84. https://doi.org/10.3354/ame032073
36. Jones AC, Liao TSV, Najar FZ, Roe BA, Hambright KD, Caron DA (2013) Seasonality and disturbance: annual pattern and response of the bacterial and microbial eukaryotic assemblages in a freshwater ecosystem. Environ. Microbiol. 15:2557–2572. https://doi.org/10.1111/1462-2920.12151
37. Amaral-Zettler LA, McCliment EA, Ducklow HW, Huse SM (2009) A method for studying protistan diversity using massively parallel sequencing of V9 hypervariable regions of small-subunit ribosomal RNA genes. PLoS One 4:e6372
38. Sogin ML, Morrison HG, Huber JA, Welch DM, Huse SM, Neal PR, Arrieta JM, Herndl GJ (2006) Microbial diversity in the deep sea and the underexplored "rare biosphere". Proc. Natl. Acad. Sci. U. S. A. 103:12115–12120
39. Huse S, Huber J, Morrison H, Sogin M, Welch D (2007) Accuracy and quality of massively parallel DNA pyrosequencing. Genome Biol. 8:R143
40. Huse SM, Welch DM, Morrison HG, Sogin ML (2010) Ironing out the wrinkles in the rare biosphere through improved OTU clustering. Environ. Microbiol. 12:1889–1898
41. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF (2009) Introducing Mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl. Environ. Microbiol. 75:7537–7541. https://doi.org/10.1128/aem.01541-09
42. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25:3389–3402
43. Pruesse E, Quast C, Knittel K, Fuchs B, Ludwig W, Peplies J, Glöckner FO (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res. 35:7188–7196
44. Clarke KR (1993) Non-parametric multivariate analyses of changes in community structure. Australian J Ecol 18: 117–143
45. Clarke KR, Warwick RM (2001) Change in marine communities: an approach to statistical analysis and interpretation, 2nd, Plymouth, UK
46. Chow C-ET, Kim DY, Sachdeva R, Caron DA, Fuhrman JA (2014) Top-down controls on bacterial community structure: microbial network analysis of bacteria, T4-like viruses and protists. ISME J 8:816–829
47. Ruan Q, Dutta D, Schwalbach MS, Steele JA, Fuhrman JA, Sun F (2006) Local similarity analysis reveals unique associations among marine bacterioplankton species and environmental factors. Bioinformatics 22:2532–2538. https://doi.org/10.1093/bioinformatics/btl417
48. Storey JD (2002) A direct approach to false discovery rates. J Royal Stat Soc: Series B (Stat Methodol) 64:479–498. https://doi.org/10.1111/1467-9868.00346
49. Xia L, Steele J, Cram J, Cardon Z, Simmons S, Vallino J, Fuhrman J, Sun F (2011) Extended local similarity analysis (eLSA) of microbial community and other time series data with replicates. BMC Systems Biol 5:S15
50. Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, Workman C, Christmas R, Avila-Campilo I, Creech M, Gross B, Hanspers K, Isserlin R, Kelley R, Killcoyne S, Lotia S, Maere S, Morris J, Ono K, Pavlovic V, Pico AR, Vailaya A, Wang P-L, Adler A, Conklin BR, Hood L, Kuiper M, Sander C, Schmulevich I, Schwikowski B, Warner GJ, Ideker T, Bader GD (2007) Integration of biological networks and gene expression data using Cytoscape. Nat. Protoc. 2:2366–2382
51. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 13:2498–2504. https://doi.org/10.1101/gr.1239303
52. Smoot ME, Ono K, Ruscheinski J, Wang P-L, Ideker T (2011) Cytoscape 2.8: new features for data integration and network visualization. Bioinformatics 27:431–432. https://doi.org/10.1093/bioinformatics/btq675
53. Assenov Y, Ramírez F, Schelhorn S-E, Lengauer T, Albrecht M (2008) Computing topological parameters of biological networks. Bioinformatics 24:282–284. https://doi.org/10.1093/bioinformatics/btm554
54. Erdös P, Réyni A (1960) On the evolution of random graphs. Inst Math Hungarian Acad Sci 5:17–61
55. Bader G, Hogue C (2003) An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinform 4:2
56. Jost L (2006) Entropy and diversity. Oikos 113:363–375
57. Friedman J, Alm EJ (2012) Inferring correlation networks from genomic survey data. PLoS Comp Biol 8(9):e1002687
58. Barabasi A-L, Oltvai ZN (2004) Network biology: understanding the cell's functional organization. Nature Rev Gen 5:101–113
59. Lima-Mendez G, van Helden J (2009) The powerful law of the power law and other myths in network biology. Molec BioSys 5:1482–1493
60. de Vargas C, Audic S, Henry N, Decelle J, Mahé F, Logares R, Lara E, Berney C, Le Bescot N, Probert I, Carmichael M, Poulain J, Romac S, Colin S, Aury J-M, Bittner L, Chaffron S, Dunthorn M, Engelen S, Flegontova O, Guidi L, Horák A, Jaillon O, Lima-Mendez G, Lukeš J, Malviya S, Morard R, Mulot M, Scalco E, Siano R, Vincent F, Zingone A, Dimier C, Picheral M, Searson S, Kandels-Lewis S, Coordinators TO, Acinas SG, Bork P, Bowler C, Gorsky G, Grimsley N, Hingamp P, Iudicone D, Not F, Ogata H, Pesant S, Raes J, Sieracki ME, Speich S, Stemmann L, Sunagawa S, Weissenbach J, Wincker P, Karsenti E (2015) Eukaryotic plankton diversity in the sunlit ocean. Science 348:1261605. https://doi.org/10.1126/science.1261605
61. Watts DJ, Strogatz SH (1998) Collective dynamics of "small-world" networks. Nature 393:440–442

62. Gobler CJ, Sunda WG (2012) Ecosystem disruptive algal blooms of the brown tide species, *Aureococcus anophagefferens* and *Aureoumbra lagunensis*. Harmful Algae 14:36–45
63. Igarashi T, Satake M, Yasumoto T (1996) Prymnesin-2: a potent ichthyotoxic and hemolytic glycoside isolated from the red tide alga *Prymnesium parvum*. J. Am. Chem. Soc. 118:479–480. https://doi.org/10.1021/ja9534112
64. Hambright KD, Beyer JE, Easton JD, Zamor RM, Easton AC, Hallidayschult TC (2015) The niche of an invasive marine microbe in a subtropical freshwater impoundment. ISME J 9:256–264. https://doi.org/10.1038/ismej.2014.103
65. Roelke DL, Barkoh A, Brooks BW, Grover JP, Hambright KD, LaClaire JW, Moeller PDR, Patino R (2015) A chronicle of a killer alga in the west: ecology, assessment, and management of *Prymnesium parvum* blooms. Hydrobiologia 764:29–50. https://doi.org/10.1007/s10750-015-2273-6
66. Acosta F, Zamor R, Najar F, Roe B, Hambright KD (2015) Dynamics of an experimental microbial invasion. Proc. Natl. Acad. Sci. U. S. A. 112:11594–11599
67. Michaloudi E, Moustaka-Gouni M, Gkelis S, Pantelidakis K (2008) Plankton community structure during an ecosystem disruptive algal bloom of *Prymnesium parvum*. J. Plankton Res. 31:301–309