SHORT COMMUNICATION

A decade of seasonal dynamics and co-occurrences within freshwater bacterioplankton communities from eutrophic Lake Mendota, WI, USA

Emily L Kara[1], Paul C Hanson[2], Yu Hen Hu[3], Luke Winslow[2] and Katherine D McMahon[1,4]

[1]Civil and Environmental Engineering Department, University of Wisconsin-Madison, Madison, WI, USA; [2]Center for Limnology, University of Wisconsin-Madison, Madison, WI, USA; [3]Department of Electrical and Computer Engineering, University of Wisconsin-Madison, Madison, WI, USA and [4]Bacteriology Department, University of Wisconsin-Madison, Madison, WI, USA

With an unprecedented decade-long time series from a temperate eutrophic lake, we analyzed bacterial and environmental co-occurrence networks to gain insight into seasonal dynamics at the community level. We found that (1) bacterial co-occurrence networks were non-random, (2) season explained the network complexity and (3) co-occurrence network complexity was negatively correlated with the underlying community diversity across different seasons. Network complexity was not related to the variance of associated environmental factors. Temperature and productivity may drive changes in diversity across seasons in temperate aquatic systems, much as they control diversity across latitude. While the implications of bacterioplankton network structure on ecosystem function are still largely unknown, network analysis, in conjunction with traditional multivariate techniques, continues to increase our understanding of bacterioplankton temporal dynamics.

The central role of seasonal cycles in the temporal dynamics of aquatic microbial communities has been highlighted in river, lake and marine systems by the use of multivariate ordinations, analysis of community similarity, diversity and the abundance of individual taxa (Crump and Hobbie, 2005; Fuhrman *et al.*, 2006; Wu and Hahn, 2006; Shade *et al.*, 2007; Gilbert *et al.*, 2012), and bacterioplankton have been linked to the seasonal succession of other planktonic organisms (Gasol *et al.*, 1992; Kent *et al.*, 2007; Maurice *et al.*, 2010). As the size and extent of biological data sets grow, scientists turn to new techniques, such as network analysis, to understand biological complexity over large scales (for example, Barberán *et al.*, 2011). For network analysis of microbial datasets, topological 'co-occurrence' networks are generated from correlative metrics, where nodes represent observed variables, and significant correlations are represented by the edges connecting them. We used an unprecedented

decade-long time series of freshwater bacterioplankton molecular community fingerprints to test the following hypotheses: (1) community co-occurrence networks from this sample set are non-random, (2) season explains the organization and complexity of co-occurrence networks and (3) community richness and diversity correlate to co-occurrence network complexity.

Surface water samples were collected from Lake Mendota, WI, USA, and analyzed by automated ribosomal intergenic spacer analysis as described previously (Shade *et al.*, 2007; see Supplementary Materials). From 2000 to 2009, a total of 34 spring, 53 summer and 34 autumn observations were made (Supplementary Table S1). Thirty-two environmental variables were collected at the same location by the North Temperate Lakes Long Term Ecological Research program (lter.limnology. wisc.edu, Supplementary Table S2).

We quantified the richness, evenness and diversity of the underlying communities, as well as the properties of spring, summer and autumn bacterioplankton community and biogeochemical co-occurrence networks. To characterize the underlying bacterial communities, richness (R, number of operational taxonomic units observed), Shannon evenness and diversity were calculated for each

season. Using local similarity analysis, co-occurrence networks (networks hereafter) were generated from correlations with local similarity scores $(R) > 0.3$, P-value < 0.001, false discovery rate < 0.05 and no time lag (Ruan et al., 2006). Erdős-Rényi random networks were generated for comparison to empirical networks, where for every empirical network with n nodes and m edges, a random network with n nodes and m edges was generated, with each edge having an equal probability of being assigned to any node. Network comparisons were made using connectance (edges (L) per node (S)^2), characteristic path length (D, average shortest path length between nodes) and clustering coefficient (Cl, of node u: $Cl(u) = 2L_u/((\deg(u)*(\deg(u) - 1))$, where L_u is the number of edges between the neighbors of u and deg is degree, the number of edges per node) (Kuchaiev et al., 2011). The properties S, L, Cl, and D were calculated for real and random networks using GraphCrunch2 software (Kuchaiev et al., 2011), and networks were visualized using Cytoscape (Smoot et al., 2011).

Network properties depended upon the number of observations included in the analysis, similar to the effect of sampling effort on the network properties described by Martinez et al. (1999). To correct for unequal sampling efforts between seasons, we performed Monte Carlo simulations of 1000 observation-normalized networks for each season, where each network was generated from 30 randomly selected observational dates within a season. For each Monte Carlo simulation, an Erdős-Rényi random graph was generated. For the networks generated from all observations from a single season, 1000 Erdős-Rényi random graphs were generated; median values for the network properties of these random graphs are reported (Table 1). Differences between median values for the distributions of randomized networks were made using the Wilcoxon rank sum test.
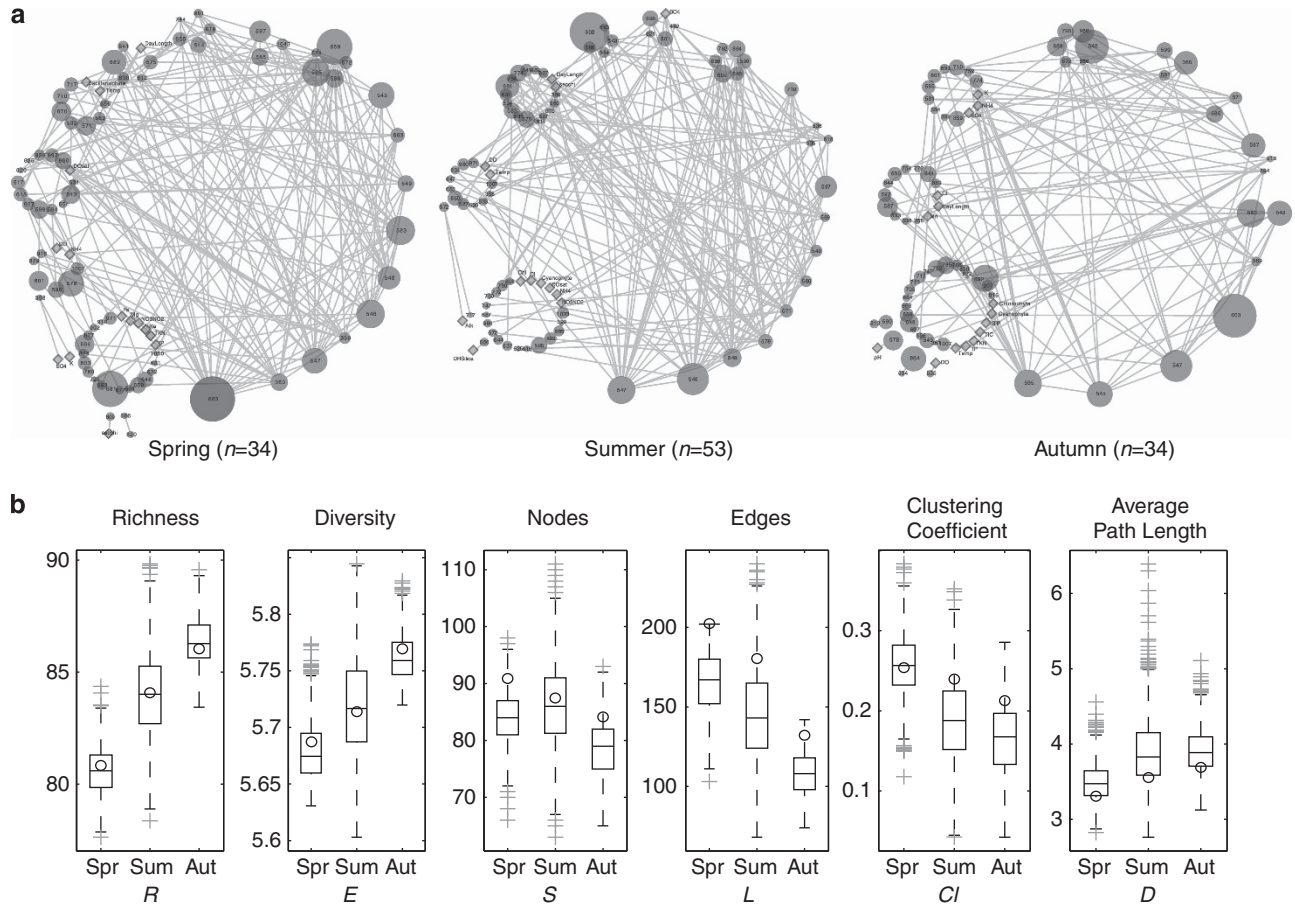
From spring to autumn, network complexity decreased, as indicated by decreasing Cl and increasing D (Figure 1a). Trends were robust to normalizing for the number of observations made: median values of observation-normalized network properties L, S, Cl and D significantly varied across season (Figures 1b, $\alpha = 0.005$), consistent with trends shown in Figure 1a. The spring network had seven high-degree nodes (> 15 edges/node), while summer and autumn networks had five and one, respectively (Figure 1a). The underlying communities' richness and diversity increased (Figure 1a) and were significantly different (Figures 1b, $\alpha = 0.005$), while community evenness did not differ significantly across seasons (data not shown). Network properties L, S, Cl and D of Monte Carlo distributions of real networks were significantly different from those of the corresponding distributions of random networks ($\alpha = 0.005$). R-values for significant pairs comprising the networks shown in Figure 1 are reported in Supplementary Materials. We examined the effect of lower ($R = 0.1$) and higher ($R = 0.5$) thresholds for the correlative local similarity score on global network properties and found that trends in complexity across seasons were generally similar to the results reported here for local similarity score $R = 0.3$ (Supplementary Figure S2).

Our results indicate: (1) non-random bacterioplankton co-occurrence networks and (2) an inverse relationship between network complexity and diversity and richness of the underlying communities for this system. Springtime conditions were favorable to fewer taxa (lower richness), but resulted in overall more co-occurrences between taxa, especially between the most abundant taxa (Figure 1). Barberán et al. (2011) hypothesized that soil microbial co-occurrence network complexity was inversely related to habitat heterogeneity across broad spatial scales. In contrast, we found that network complexity was not related to heterogeneity of community habitat, as estimated by the variance of

Table 1 Summary of co-occurrence network properties

	Nodes (S)	Edges (L)	L/S	L/S^2	D	D$_{rand}$	Cl	Cl$_{rand}$	Cl/Cl$_{rand}$
All observations									
Spring	91	202	2.22	0.0244	3.27	**3.41**	0.248	**0.045**	5.57
Summer	88	182	2.07	0.0235	3.54	**3.83**	0.236	**0.043**	5.55
Autumn	84	131	1.56	0.0186	3.59	**3.82**	0.217	**0.029**	7.43
Sampling effort-normalized median values from Monte Carlo simulations									
Spring	**84**	**167**	**1.99**	0.0237	3.47	3.66	0.256	0.043	5.90
Summer	**86**	**143**	**1.66**	0.0193	3.83	4.13	0.188	0.032	5.81
Autumn	**79**	**108**	**1.37**	0.0173	3.88	4.18	0.167	0.026	6.56

Regular type indicates a single value calculated and reported, while boldface type indicates the median value reported from an analysis including 1000 Monte Carlo simulations. Top panel: network properties calculated from networks generated from all observations (n) made for spring ($n = 34$), summer ($n = 53$) and autumn ($n = 34$) seasons. For networks generated from all observations, 1000 Erdős-Rényi (E-R) random graphs were generated with an equal number of nodes (S) and edges (L) as the corresponding original network, but distinct characteristic path length (D) and clustering coefficient (Cl); we report the median values of random graph populations, indicated in bold (D_{rand} and Cl_{rand}). Bottom panel: network properties from observation-normalized Monte Carlo simulations ($n = 30 \times 1000$ simulations) and Erdős-Rényi random graphs generated from each Monte Carlo simulation. As above, random graphs had equal L and S but significant differences between D and D_{rand} and Cl and Cl_{rand}.

Figure 1 Co-occurrence network visualization (**a**) and properties (**b**) for a 10-year time series of bacterioplankton communities in Lake Mendota, WI, USA. From spring to autumn, diversity (*E*) and richness (*S*) increased while network complexity decreased. The number of total observations (*n*) for each season are indicated in parentheses below the networks. Nodes are labeled with operational taxonomic unit (OTU) fragment length or environmental variable name. For operational taxonomic units, node area is proportional to average relative abundance for the season indicated. Diamonds indicate all environmental variable nodes. Nodes are grouped according to degree (edges/node), with the highest degree node at the bottom of the network and nodes of decreasing degree placed sequentially, counter-clockwise, such that lowest-degree nodes (1 edge/node) are grouped to the immediate left of the highest-degree node. Network complexity decreased across seasons while richness and diversity increased, and trends in all network properties, diversity and richness were robust to normalizing for the number of observations made (**b**). Network properties (number of nodes (*S*), edges (*L*), clustering coefficient (Cl), and characteristic path length (*D*)) are shown in (**b**). Open circles indicate properties calculated for networks shown in (**a**), which included all observations for a given season. Boxplots represent observation-normalized results, where boxes represent inter-quartile range, whiskers indicate 10th and 90th percentiles and plus symbols indicate outliers.

32 physical, chemical and biological variables across seasons, or to the complexity of co-occurrence networks of environmental variables for each season (Supplementary Figure S1).

Correlations (edges) between temperature, day length, dissolved oxygen and bacterial taxa were more frequent than those between taxa and other variables (Figure 1a and local similarity results reported in Supplementary Materials). Increases in bacterioplankton community diversity and richness over the open water season may be due to changes in day length, and concomitant trends in water temperature and primary productivity, which lags insolation by ~7 weeks in this system (Brock, 1985). Diversity and latitude are negatively related in marine bacterial communities (Pommier *et al.*, 2007; Fuhrman *et al.*, 2008) and putative causes of

the relationship are: (1) higher productivity at lower absolute latitudes accommodates greater diversity (Connell and Orias, 1964) and (2) temperature controls biological process rates, including speciation (Rohde, 1992). We hypothesize that in temperate regions, annual cycles in temperature and productivity control the diversity of aquatic microbial communities through time, much like latitude controls diversity globally. Our results indicate that these variables have an additive effect on diversity, such that cumulative productivity after ice-off controls diversity. Annually, communities seem to 'reset' under ice, where diversity and co-occurrence network structure converge to characteristic spring-time conditions (lower diversity and higher co-occurrence network complexity) due to low temperatures and greatly reduced insolation

beneath ice cover. Uneven temporal sampling within this dataset limits our ability to test these hypotheses, but cross-system comparisons over a gradient of lake trophic status and latitude could be used to do so.

The robust recurrent patterns in diversity and co-occurrence networks are especially compelling in light of the magnitude of microbial turnover rates (10–60% removal of standing bacterial biomass per day) measured in this system and others (Brock, 1985; Hahn et al., 2012). Although a vast and diverse persistent seed bank likely underlies the observed communities (Caporaso et al., 2011; Lennon and Jones, 2011), the most abundant taxa co-occur in a repeatable pattern despite constant removal by losses due to planktonic grazers or viral lysis. The repeatable co-occurrence of members of the community may be determined by environmental cues that trigger growth from dormant low abundance when more favorable conditions occur (Jones and Lennon, 2010).

In this early stage of the application of network analysis to microbial observations, the implications of the co-occurrence network structure on the stability and function of ecosystems and engineered systems are unknown. In macro-scale ecology, communities with high diversity and low co-occurrence, like those represented by the autumnal networks, have been observed and associated with greater stability (McCann, 2000). In contrast, food web stability has been shown to correspond to higher complexity networks (Dunne et al., 2002). The possibilities of testing stability–network complexity relationships in microbial systems that are easily or stochastically perturbed (for example, microbially-mediated wastewater treatment processes or human gut microbial communities) are numerous, and, due to the availability of economical and high-throughput next-generation sequencing technologies, are within reach for experimental ecology.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

References

Barberán A, Bates ST, Casamayor EO, Fierer N. (2011). Using network analysis to explore co-occurrence patterns in soil microbial communities. ISME J 6: 343–351.

Brock TD. (1985). A Eutrophic Lake: Lake Mendota, Wisconsin. Springer Verlag: New York, p 308.

Caporaso JG, Paszkiewicz K, Field D, Knight R, Gilbert JA. (2011). The Western English Channel contains a persistent microbial seed bank. ISME J 6: 1089–1093.

Connell JH, Orias E. (1964). The ecological regulation of species diversity. Am Nat 98: 399–414.

Crump BC, Hobbie JE. (2005). Synchrony and seasonality in bacterioplankton communities of two temperate rivers. Limnol Oceanogr 50: 1718–1729.

Dunne JA, Williams RJ, Martinez ND, Fe S. (2002). Network topology and biodiversity loss in food webs: Robustness increases with connectance. Ecol Lett 5: 558–567.

Fuhrman JA, Hewson I, Schwalbach MS, Steele JA, Brown MV, Naeem S. (2006). Annually reoccurring bacterial communities are predictable from ocean conditions. Proc Natl Acad Sci USA 103: 13104–13109.

Fuhrman JA, Steele JA, Hewson I, Schwalbach MS, Brown MV, Green JL et al. (2008). A latitudinal diversity gradient in planktonic marine bacteria. Proc Natl Acad Sci Usa 105: 7774–7778.

Gasol J, Peters F, Guerrero R, Pedrós-Alió C. (1992). Community structure in Lake Ciso: Biomass allocation to trophic groups and differing patterns of seasonal succession in the meta- and epilimnion. Arch Hydrobiol 123: 275–303.

Gilbert JA, Steele JA, Caporaso JG, Steinbrück L, Reeder J, Temperton B et al. (2012). Defining seasonal marine microbial community dynamics. ISME J 6: 298–308.

Hahn MW, Scheuerl T, Jezberová J, Koll U, Jezbera J, Šimek K et al. (2012). The passive yet successful way of planktonic life: genomic and experimental analysis of the ecology of a free-living polynucleobacter population. PLoS One 7: e32772: 1-17.

Jones SE, Lennon JT. (2010). Dormancy contributes to the maintenance of microbial diversity. Proc Natl Acad Sci Usa 107: 5881–5886.

Kent AD, Yannarell AC, Rusak JA, Triplett EW, Mcmahon KD. (2007). Synchrony in aquatic microbial community dynamics. ISME J 1: 38–47.

Kuchaiev O, Stevanović A, Hayes W, Pržulj N. (2011). GraphCrunch 2: Software tool for network modeling, alignment and clustering. BMC bioinformatics 12: 24.

Lennon JT, Jones SE. (2011). Microbial seed banks: the ecological and evolutionary implications of dormancy. Nat Rev Microbiol 9: 119–130.

Martinez ND, Hawkins BA, Dawah HA, Feifarek BP. (1999). Effects of sampling effort on characterization of food-web structure. Ecology 80: 1044–1055.

Maurice CF, Bouvier T, Comte J, Guillemette F, Del Giorgio PA. (2010). Seasonal variations of phage life strategies and bacterial physiological states in three northern temperate lakes. Environ Microbiol 12: 628–641.

McCann KS. (2000). The diversity-stability debate. Nature 405: 228–233.

Pommier T, Canbäck B, Riemann L, Boström KH, Simu K, Lundberg P et al. (2007). Global patterns of diversity and community structure in marine bacterioplankton. Mol Ecol 16: 867–880.

Rohde K. (1992). Latitudinal gradients in species the search for the diversity: the search for primary cause. *Oikos* **65**: 514–527.

Ruan Q, Steele JA, Schwalbach MS, Fuhrman JA, Sun F. (2006). A dynamic programming algorithm for binning microbial community profiles. *Bioinformatics (Oxford, England)* **22**: 1508–1514.

Shade A, Kent AD, Jones SE, Newton RJ, Triplett EW, McMahon KD. (2007). Interannual dynamics and phenology of bacterial communities in a eutrophic lake. *Limnol Oceanogr* **52**: 487–494.

Smoot M, Ono K, Ruscheinski J, Wang P, Ideker T. (2011). Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics (Oxford, England)* **27**: 431–432.

Wu QL, Hahn MW. (2006). High predictability of the seasonal dynamics of a species-like Polynucleobacter population in a freshwater lake. *Environ Microbiol* **8**: 1660–1666.

Supplementary Information accompanies the paper on The ISME Journal website (http://www.nature.com/ismej)