USING ECOLOGICAL THEORY TO INVESTIGATE EMERGENT PROPERTIES OF

POPULATIONS IN AQUATIC ECOSYSTEMS


By

Cristina Marie Herren


A dissertation submitted in partial fulfillment of

the requirements for the degree of


Doctor of Philosophy

(Freshwater and Marine Science)


at the

UNIVERSITY OF WISCONSIN-MADISON

2016

Date of final oral examination: 20 June 2017

The dissertation is approved by the following members of the Final Oral Committee:
    Dr. Katherine D. McMahon, Professor, Civil and Environmental Engineering & Bacteriology
    Dr. Claudio Gratton, Professor, Entomology
    Dr. Emily H. Stanley, Professor, Zoology & Center for Limnology
    Dr. M. Jake Vander Zanden, Professor, Zoology & Center for Limnology
    Dr. Paul C. Hanson, Distinguished Professor of Research, Center for Limnology

**Abstract**

Populations behave inherently differently than individuals. The features that arise when individuals aggregate and interact, such as population oscillations and stable age distributions, are called emergent properties. Ecologists have studies these properties for decades, especially when they pertain to sudden, dramatic shifts in population size. However, empirical studies are less common, because it is difficult to meet the assumptions of theoretical models in real systems. This dissertation applies ecological theories to several different aquatic systems to better understand and model characteristics of these ecosystems, many of which are the results of emergent properties. Chapter 2 examines how environmental disturbances affect the variability of diatom and bacteria populations within biofilms. I found that experimentally induced environmental stressors acted as deterministic, selective forces in these communities, thereby creating populations that were more similar to one another after being disturbed. Chapter 3 was prompted by the observation that the primary and secondary productivity of Lake Myvatn, a sub-arctic lake in northeast Iceland, were extremely high, given its latitude. I hypothesized that the secondary producers, which are predominantly midges, were involved in a mutualism that enabled high growth rates of both algae and midge larvae. This study found that the midges were able to alleviate their own resource limitation by promoting the growth of their benthic algal resources, thereby increasing both primary and secondary production. Chapters 4 and 5 are paired chapters that develop a novel statistical workflow (Chapter 4) and implement this analysis on a variety of long-term microbial datasets (Chapter 5). One of the earliest questions in theoretical ecology asked how the complexity of food webs related to the stability of these systems. This question is often intractable due to the need to observe hundreds of taxa over many generations, but bacterial systems overcome this challenge. In Chapter 4, I address this question

by creating a method to quantify the connectedness of ecological communities, which is one aspect of community complexity. In Chapter 5, I applied this workflow to three long-term microbial datasets, and found that highly connected keystone taxa have disproportionate influence in predicting compositional turnover in the entire community.

**Dedication**

During the course of my scientific career, I have been fortunate to receive support and guidance from many people. In chronological order, I would like to thank those who have contributed to my progress as a researcher, teacher, and mentor. First, I would like to thank my high school biology teacher, Ms. Lidia Ortiz, for encouraging me to pursue science at a time when I had not yet recognized my passion for it. I am also grateful to Dr. Matt Ayres, who taught my first ecology course at Dartmouth College, which was so engaging that I quickly enrolled in more classes. Matt also directed me to Dr. Mark McPeek, who had a profound influence on my undergraduate science career as my senior thesis advisor. I am particularly thankful to Mark for teaching scientific ethics as a part of his classes. Furthermore, I am indebted to Dr. Kathryn Cottingham for her dedicated mentoring and for giving me a first exposure to scientific research, which led me to apply to graduate school.

While at UW-Madison, I have had the pleasure of knowing many outstanding scientists and friends. I sincerely thank my advisor, Dr. Trina McMahon, for her support and perspective. My committee members, including Dr. Emily Stanley, Dr. Claudio Gratton, Dr. Jake Vander Zanden, and Dr. Paul Hanson, were invaluable resources. I also have deep gratitude for my friends and labmates, who have undoubtedly contributed to my personal and scientific successes. In particular, I thank Sarah Stevens and Dr. Kyle Webert for their unwavering encouragement and confidence. I am also especially grateful to other UW-Madison graduate students, including Dr. Rose Graves, Hilary Barker, and Rachael Zinn. Finally, my friends from near and far have been an exceptional support network, and I would like to thank Dr. Cayelan Carey, Dr. Laurel Symes, Dr. Jessica Trout-Haney, Carol Warden, Susan Hakes, Margaret Jodlowski, Sally Batton, Priya Shanmugam, Leigh Latimer, Giulia Siccardo, Kevin Davies, and Mark Bridges.

**Table of Contents**

**Introduction**

Complex dynamics can arise even in simple systems (Lorenz 1963, May and Oster 1976).

One of the distinguishing features of population and community ecology is the study of emergent

properties, which arise from groups of interacting individuals (Odum and Barrett 1971). Many of

these emergent properties at the population level pertain to population variability, such as the

properties of population regulation or population oscillations (Berryman et al. 2002).

Furthermore, when multiple populations interact to form ecological communities, additional

phenomena are possible; although there are many emergent properties of ecological

communities, one of the most well studied is community stability (May 1973). This concept,

along with its implied notion of an equilibrium community, has been central to community

ecology for over a century (Clements 1916, Whittaker 1953, MacArthur 1955, May 1972,

McCann 2000). Although emergent properties are understood well through mathematical

models, it is often difficult to test predictions of these models empirically, because few systems

meet the assumptions of the theoretical models.

This dissertation applies ecological theory to aquatic communities to investigate the

emergent properties within these ecosystems. Chapters 2 and 3 explore causes of population

variability and mechanisms contributing to population regulation. Specifically, Chapter 2 asks

whether environmental stressors, acting as forces of natural selection, can alter the inherent level

of variability present in natural populations. Chapter 3 investigates whether the positive

interactions between two trophic levels might alter ecosystem productivity and increase

population size of the species involved in the positive feedback. Chapters 4 and 5 are paired

chapters that empirically investigate a decades-old hypothesis about how the interconnectedness

of ecological communities relates to their rate of compositional change. Chapter 4 develops and

validates the workflow for this analysis, while Chapter 5 tests the hypothesis on several long-term datasets and two case studies.

Chapter 2 uses the framework of deterministic versus stochastic processes (Vellend 2010) to investigate whether environmental disturbances alter the amount of variability present in populations. All natural populations have inherent variability, both because they exist in dynamic environmental conditions (Lundberg et al. 2000) and because there is some degree of chance in forces such as dispersal and resource patchiness (Nemergut et al. 2013). This chapter asked whether environmental stress could alter population variability, because selection is presumed to be a deterministic force (Nemergut et al. 2013). During the summer of 2013, I conducted an experiment where I grew many biofilm communities under similar environmental conditions. Then, these "replicate" communities were randomly divided into nine treatments, each of which experienced a different disturbance regime. At the end of the experiment, I sampled both the diatom communities (through direct cell counts under a microscope) and the bacterial communities (using a fingerprinting technique that gave relative abundances of operational taxonomic units [OTUs]). I compared the variability of populations within the disturbed communities to the variability of populations in the undisturbed communities. In all instances where the treatment effect was significant, the experimental disturbances decreased population variability. Thus, this chapter suggests that environmental stress can create populations that are more predictable, in the sense that their composition might be more accurately forecasted due to lower population variability. Additional work is needed to uncover the mechanisms driving these patterns, which is discussed as Project 2 in Chapter 6.

Chapter 3 began with the observation that the chironomid larvae of Lake Myvatn experienced exponential growth for a surprisingly long duration, increasing by an order of

magnitude for 5-6 consecutive generations (Einarsson et al. 2002). This observation indicated that the chironomid population was not experiencing population regulation during this exponential growth phase, which suggested that resources were not limiting to the population (Barryman et al. 2002). I hypothesized that the chironomid larvae, which are known to be ecosystem engineers (Holker et al. 2015), were acting as "farmers" and alleviating their own resource limitation by promoting algal growth. Thus, the ecological theory guiding this chapter was the effect of mutualisms on equilibrium population densities and realized population growth rates. The results of experiments in this chapter showed that chironomid larvae strongly increased both standing algal biomass and rates of primary production. The positive feedback between chironomid larvae and benthic algae was sufficiently strong that chironomids grew more quickly when stocked at higher initial densities. Thus, this empirical result supports the theoretical hypothesis that mutualisms can increase densities of both populations involved in the mutualism (Holland and DeAngelis 2010).

The next two chapters, 4 and 5, developed from my longstanding fascination with using coupled differential equations to model ecological communities. One of the most famous empirical examples to validate this type of model was the demonstration of trophic cascades (Carpenter et al. 1985). As predicted from the equations, adding a top predator can alter the population sizes of species within lower trophic levels. The topic in this realm of community ecology that I have always found most interesting is how the complexity of communities relates to their stability. Historically, this question has been investigated by using linear stability analysis with differential equations that express the various growth rates of, and interactions between, species (May 1973). However, this approach has several limitations, as some of these mathematical quantities have no clear analog in real systems (i.e. the definition of stability as the

sign of the dominant eigenvalue of the Jacobian matrix). Furthermore, any study that seeks to model the connectedness of natural communities must analyze datasets that meet the following criteria: 1) a sufficiently long time series to span the range of environmental conditions several times 2) enough taxa to be able to characterize statistical properties of taxon populations 3) enough samples to have good resolution when inferring interactions between taxa. I believed that bacteria might be a good system to investigate this topic, because bacterial communities overcame these major initial hurdles. The question I asked was similar to the theoretical postulation, but not exactly the same; instead, I asked how the interconnectedness of microbial communities relates to their rate of compositional turnover through time. However, bacterial datasets have their own drawbacks (such as often being in the form of relativized data and being strongly zero-inflated), so I also invested a substantial amount of time in creating, testing, and validating the workflow for this analysis. The two components of this project, designing the workflow and analyzing several long-term datasets, became Chapters 4 and 5.

Ecological theory suggests that the connectivity of communities should be related to their emergent properties, including their stability. However, quantifying this connectivity is difficult in real systems. Chapter 4 develops a method for calculating the connectedness of microbial taxa, which is then used to quantify the connectivity of microbial communities. A major strength of this workflow, as compared to other existing methods of analyzing community connectedness, is that it yields two values of connectivity for each sampled community. In contrast, other methods often result in large network models with many parameters, which are difficult to interpret. The benefit of having values of connectivity ascribed to each sample is that this workflow can be integrated with other analyses. For example, this chapter demonstrates how the values of connectivity can be used as predictors in multiple regressions. Thus, this workflow is compatible

with many other existing ways to analyze ecological data. To demonstrate the utility of this approach, this chapter uses 5 long-term datasets to show that the workflow shows a significant relationship between community connectivity and Bray-Curtis dissimilarity of phytoplankton communities. Finally, this chapter validates each step of the workflow to assess and account for potential pitfalls of working with microbial datasets.

Recently, microbial ecologists have been applying the concept of keystone species to microbial communities to identify taxa that have a disproportionate influence in community dynamics (Vick-Majors et al. 2014, Agler et al. 2016). One proposed method for identifying keystone taxa is looking for taxa that are highly connected hubs in the milieu of microbial interactions (Banerjee et al. 2016). Chapter 5 uses the workflow I designed in Chapter 4 to first test whether my metrics of community cohesiveness relate to the rate of compositional turnover, and then analyze whether the highly connected taxa are most informative about pending compositional changes. I had hypothesized that a small subset of keystone taxa were responsible for driving community dynamics. I used three long-term microbial datasets to analyze how the cohesiveness of microbial communities related to their rate of compositional change, and found in each case that my metrics of cohesion were strong predictors of Bray-Curtis dissimilarity; communities that were more highly connected, especially if the connectivity arose from negative correlations between taxa, were much more compositionally stable through time. Surprisingly, these models, which included no environmental parameters, sometimes outperformed other analyses of these datasets that used dozens of supporting environmental datasets. Furthermore, I found that including only the most highly connected taxa (1-5% of richness) produced models that were as informative or more informative than including all taxa. Thus, these keystone microbes were particularly revealing about whole-community properties. One main practical

goal of this chapter was to identify the taxa that microbial ecologists should focus on in pursuit of understanding community dynamics. The high diversity of these communities often renders it overwhelming to choose individual taxa to study, but I am optimistic that this analysis might help direct researchers to taxa that have disproportionate importance in the community.

The thread connecting these chapters is the use of conceptual models and ecological theory to investigate the distinguishing characteristics of empirical ecological communities. This work is based upon the premise that emergent properties only arise when particular mechanisms are operating in populations or communities. Thus, studying the connection between mechanisms and emergent properties can yield insights when beginning from either direction; knowing what mechanisms are present in a system lends predictive power to the long-term emergent dynamics of a community, and observing emergent properties of a community suggests that certain mechanisms are operating in that system.

**References**

Agler, M. T., J. Ruhe, S. Kroll, C. Morhenn, S.-T. Kim, D. Weigel, and E. M. Kemen. 2016. Microbial Hub Taxa Link Host and Abiotic Factors to Plant Microbiome Variation. PLOS Biol 14:e1002352.

Banerjee, S., C. A. Kirkby, D. Schmutter, A. Bissett, J. A. Kirkegaard, and A. E. Richardson. 2016. Network analysis reveals functional redundancy and keystone taxa amongst bacterial and fungal communities during organic matter decomposition in an arable soil. Soil Biology and Biochemistry 97:188–198.

Berryman, A. A., M. Lima Arce, and B. A. Hawkins. 2002. Population regulation, emergent properties, and a requiem for density dependence. Oikos 99:600–606.

Carpenter, S. R., J. F. Kitchell, and J. R. Hodgson. 1985. Cascading Trophic Interactions and Lake Productivity. BioScience 35:634–639.

Clements, F. E. 1916. Plant Succession: An Analysis of the Development of Vegetation. Carnegie Institution of Washington.

Einarsson, Á., A. Gardarsson, G. M. Gíslason, and A. R. Ives. 2002. Consumer–resource interactions and cyclic population dynamics of Tanytarsus gracilentus (Diptera: Chironomidae). Journal of Animal Ecology 71:832–845.

Hölker, F., M. J. Vanni, J. J. Kuiper, C. Meile, H.-P. Grossart, P. Stief, R. Adrian, A. Lorke, O. Dellwig, A. Brand, M. Hupfer, W. M. Mooij, G. Nützmann, and J. Lewandowski. 2015. Tube-dwelling invertebrates: tiny ecosystem engineers have large effects in lake ecosystems. Ecological Monographs 85:333–351.

Holland, J. N., and D. L. DeAngelis. 2010. A consumer–resource approach to the density-dependent population dynamics of mutualism. Ecology 91:1286–1295.

Lorenz, E. N. 1963. Deterministic Nonperiodic Flow. Journal of the Atmospheric Sciences 20:130–141.

Lundberg, P., E. Ranta, J. Ripa, and V. Kaitala. 2000. Population variability in space and time. Trends in Ecology & Evolution 15:460–464.

MacArthur, R. 1955. Fluctuations of Animal Populations and a Measure of Community Stability. Ecology 36:533–536.

May, R. M. 1972. Will a Large Complex System be Stable? Nature 238:413–414.

May, R. M. 1973. Stability and Complexity in Model Ecosystems. Princeton University Press.

May, R. M., and G. F. Oster. 1976. Bifurcations and Dynamic Complexity in Simple Ecological Models. The American Naturalist 110:573–599.

McCann, K. S. 2000. The diversity–stability debate. Nature 405:228–233.

Nemergut, D. R., S. K. Schmidt, T. Fukami, S. P. O'Neill, T. M. Bilinski, L. F. Stanish, J. E. Knelman, J. L. Darcy, R. C. Lynch, P. Wickey, and S. Ferrenberg. 2013. Patterns and Processes of Microbial Community Assembly. Microbiology and Molecular Biology Reviews 77:342–356.

Odum, E. P. and G. W. Barrett. 1971. *Fundamentals of ecology* (Vol. 3). Philadelphia: Saunders.

Vellend, M. 2010. Conceptual Synthesis in Community Ecology. The Quarterly Review of Biology 85:183–206.

Vick-Majors, T. J., J. C. Priscu, and L. A. Amaral-Zettler. 2014. Modular community structure suggests metabolic plasticity during the transition to polar night in ice-covered Antarctic lakes. The ISME Journal 8:778–789.

Whittaker, R. H. 1953. A Consideration of Climax Theory: The Climax as a Population and Pattern. Ecological Monographs 23:41–78.

**Chapter 2:** Environmental disturbances decrease the variability of microbial populations within periphyton

**Authors:** Cristina M. Herren[1], Kyle C. Webert[2], and Katherine D. McMahon[3]

**Author affiliations:**

[1]Freshwater and Marine Sciences Program, University of Wisconsin - Madison, Madison, Wisconsin, USA

cherren@wisc.edu

[2]Department of Zoology, University of Wisconsin - Madison, Madison, Wisconsin, USA

[3]Departments of Bacteriology and Civil and Environmental Engineering, University of Wisconsin - Madison, Madison, Wisconsin, USA

**Abstract**

A central pursuit of microbial ecology is to accurately model changes in microbial community composition in response to environmental factors. This goal requires a thorough understanding of the drivers of variability in microbial populations. However, most microbial ecology studies focus on the effects of environmental factors on mean population abundances, rather than on population variability. Here, we imposed several experimental disturbances upon periphyton communities and analyzed the variability of populations within disturbed communities versus those in undisturbed communities. We analyzed both the bacterial and the diatom communities in the periphyton under nine different disturbance regimes, including regimes that contained multiple disturbances. We found several similarities in the responses of the two communities to disturbance; all significant treatment effects showed that populations became less variable as the result of environmental disturbances. Furthermore, multiple disturbances to these communities were often interactive, meaning that the effects of two disturbances could not have been predicted from studying single disturbances in isolation. These results suggest that environmental factors had repeatable effects on populations within microbial communities, thereby creating communities that were more similar as a result of disturbances. These experiments add to the predictive framework of microbial ecology by quantifying variability in microbial populations and by demonstrating that disturbances can place consistent constraints on the abundance of microbial populations. Although models will never be fully predictive due to stochastic forces, these results indicate that environmental stressors may increase the ability of models to capture microbial community dynamics because of their consistent effects on microbial populations.

**Importance**

There are many reasons why microbial community composition is difficult to model. For example, the high diversity and rapid rate of change of these communities make it challenging to identify causes of community turnover. Furthermore, the processes that shape community composition can be either deterministic, which cause communities to converge upon similar compositions, or stochastic, which increase variability in community composition. However, modeling microbial community composition is only possible if microbes show repeatable responses to extrinsic forcing. In this study, we hypothesized that environmental stress acts as a deterministic force that shapes microbial community composition. Other studies have investigated if disturbances can alter microbial community composition, but relatively few studies that ask about the repeatability of the effects of disturbances. Mechanistic models implicitly assume that communities show consistent responses to stressors; here, we define and quantify microbial variability to test this assumption.

## Introduction

Many central questions in ecology focus on the sources of variability in populations. Accuracy of predictions is highly valued in ecological studies, and population size is necessarily more predictable when populations are less variable (1, 2). Still, most ecological studies measure changes in the mean number of individuals within populations, rather than the variability of populations across time or space (3, 4). However, the variability of ecological communities is sensitive to environmental drivers, and is therefore expected to change in response to disturbances (5, 6). These responses can be observed both temporally, where the variance of a population is calculated over time (7), or spatially, across a landscape or between communities (8). Here, we analyze the variability of populations between replicated microbial communities after a series of experimental disturbances. Specifically, we ask whether this strong environmental forcing creates communities where taxon abundance is less variable than in undisturbed communities. Thus, we address whether disturbances have repeatable effects on ecological communities.

Some variability naturally exists in all populations. Disturbances could either act to increase or decrease this level of variability (6, 7). The effect of the disturbance on population variability is dependent upon whether the disturbance acts as a deterministic or a stochastic force (9). For example, disturbance could act as a deterministic force to decrease variability by imposing a consistent selective pressure, creating communities that are more similar to one another (10). Conversely, disturbance could disrupt feedback loops formed by species interactions (11) and cause initially-similar communities to exhibit increased stochasticity. Under differing circumstances, both of these responses have been observed in microbial systems. For

example, bacterial communities within bioreactors showed variability in composition after being disturbed with glucose additions (12). Although there were consistent functional changes in the bioreactors, there was low replicability in bacterial community composition among reactors. However, disturbances can canalize community composition under other circumstances. Rolke et al. (13) demonstrated that nutrient pulses generated predictable succession in phytoplankton communities, whereas undisturbed communities diverged along chaotic compositional trajectories. However, because experiments studying the variability of microbial communities often use different disturbances and metrics of variation, it is difficult to draw general conclusions about the effect of environmental stress on community variability.

Predicting the composition of microbial communities using environmental disturbances is a major objective of microbial ecology (14). Here, we define a disturbance as an external force that perturbs ecological communities in such a way that it selectively favors or disfavors specific populations or interferes with community processes (15). Several studies have stated that their goal was to understand how environmental disturbances change microbial community dynamics (16-19). However, this prediction is only possible if microbial responses to environmental forcing are repeatable. Thus, to predict microbial community responses, it is first necessary to understand how environmental drivers contribute to community variability (16). Therefore, the relationship between environmental disturbances and population-level variability is important to achieving applied goals, such as modeling microbial community composition.

To address the strong interest in understanding drivers of variability in microbial community composition, some studies have analyzed the range of compositions observed in bacterial communities following novel disturbances. These experiments have found consistent changes in community composition due to the disturbance (20, 21). Thus, strong environmental

forcing induced a reproducible shift in bacterial community composition. However, experiments that analyze microbial population variability have only been conducted in a few systems, and results are often qualitative. Furthermore, these results are difficult to generalize because different ecological communities may show varied responses to the same environmental forcing (22, 23). For instance, resilient communities are characterized by a quick recovery time (24, 25), so the effects of disturbance on highly resilient communities may only be apparent for a brief period. The response to multiple disturbances is even more difficult to predict, because there are often interactive and unexpected effects of the compounded stressors (26, 27). Recognizing these challenges, our experiments were designed specifically to analyze responses of two communities experiencing the same disturbances and to examine the effects of multiple disturbances.

In this study, we imposed several disturbance regimes upon periphyton communities in order to examine the effects of disturbances on the variability of populations within the periphyton. Our goal was to determine whether disturbances have repeatable effects on periphyton communities. After initially growing 108 periphyton communities on Plexiglass slides in a common environment, we then randomized each of these replicate communities into one of nine treatments, each corresponding to a different disturbance regime. To generate these nine treatments, we subjected periphyton communities to one of three possible conditions (water scouring disturbance, altered depth in the water column disturbance, or no disturbance) at two time points. These two disturbances were selected because they are both potential consequences of the high wind events that occur in our study system (28). This 3x3 factorial design generated treatments that included different numbers of disturbance (0, 1, or 2 disturbance events) and different combinations of disturbances.

We quantified the variability of populations between communities within each treatment using the coefficient of variation (CV) of each taxon. The coefficient of variation is calculated as the standard deviation of the populations divided by the mean abundance of the populations, and therefore has the advantage of accounting for variance-mean scaling (6). We transformed or detrended CVs as necessary to ensure that this metric was approximately normally distributed and was not biased by mean population size. Then, we used linear mixed models to compare the variability of taxa in undisturbed treatments to the variability of taxa that experienced disturbance. We separately analyzed the diatoms and bacteria found within the periphyton to compare the effects of the same disturbance regimes on these two different ecological communities.

**Results**

Periphyton colonized Plexiglas slides suspended in a shallow (maximum natural depth ~4m, [28]) eutrophic lake over a period of 20 days. Experimental disturbances were then imposed at two time points (T1 and T2, corresponding to day 20 and day 25). At these two time points, communities were randomly assigned to conditions where they were either left undisturbed (Ambient), disturbed by relocating the communities at a different depth in the water column (Depth), or disturbed with water scouring (Scoured). In the altered depth disturbance, we suspended the Plexiglas slides at 0.5m depth, rather than 3m depth, for 5 days. In the water scouring manipulation, we dragged the Plexiglas slides through the water column for 10 minutes at a rate of 20-25cm/s. Both the Ambient slides and the Scoured slides were then replaced in the lake at 3m depth for 5 days. The combination of these three conditions at the first time point and three conditions at the second time point created nine treatments: Ambient/Ambient, Ambient/Depth Change, Ambient/Scoured, Depth Change/Ambient, Depth Change/Depth

Change, Depth Change/Scoured, Scoured/Ambient, Scoured/Depth Change, and

Scoured/Scoured. We henceforth refer to these treatments as AA, AD, AS, DA, DD, DS, SA,

SD, and SS.

*Diatom Communities*

Diatom taxa on slides were enumerated by light microscopy to measure the abundance of

each taxon within the periphyton biofilm. We used linear mixed models to analyze how the

variability of taxa in the disturbed treatments compared to the variability of taxa in the

undisturbed treatment, AA. The fixed effects in the model were 4 binary variables corresponding

to whether the community received the depth disturbance at T1, the scouring disturbance at T1,

the depth disturbance at T2, or the scouring disturbance at T2. A random effect for taxon was

included, under the assumption that taxa have differing amounts of inherent population

variability. The response variable in the linear mixed models was the square root CV of the taxon

populations, measured in density per $cm^2$. A lower CV corresponded to lower variability of a

taxon between communities in the same treatment. The estimated treatment effects from the

mixed model represent the mean difference in square root CVs between a given treatment and

the undisturbed treatment, AA.

The treatment that had the highest mean taxon square root CV was the undisturbed

treatment, AA. Thus, taxon populations in the AA treatment were most variable of any treatment.

The four treatments that received one disturbance, SA, AS, DA, and AD, all had significantly

lower square root CVs than the AA treatment (Table 1). These treatment estimates and the

corresponding p values were obtained from the linear mixed model for the diatom taxa. The

scouring disturbance reduced the square root CV of the diatom communities by a mean of 0.155

at Time 1 (p = 0.0343) and by 0.158 at Time 2 (p = 0.0307). The altered depth disturbance

reduced the square root CV by 0.156 at Time 1 (p = 0.0330) and by 0.230 at Time 2 (p =

0.0016). Thus, densities of diatom taxa became more consistent as a result of experiencing one

disturbance, regardless of whether the disturbance was applied at Time 1 or Time 2.

In both the DD and SS treatments, there were significant positive interactions between

the disturbances at Time 1 and Time 2 (p = 0.0458 and p = 0.0363, respectively). In these

treatments, which received the same type of disturbance at Time 1 and Time 2, the taxa were

more variable than would be expected from the independent effects of the disturbances at Time 1

and Time 2. However, there was no significant interaction between disturbances in communities

that experienced different types of disturbances at Time 1 and Time 2 (corresponding to

treatments DS and SD). Therefore, communities that received different disturbances at the two

time points continued to become less variable as a result of experiencing another disturbance,

whereas communities that experienced the same disturbance twice did not become as consistent.

We performed principal components analyses (PCAs) to determine if the community

composition shifted as a result of the disturbances. We compared the undisturbed treatment, AA,

to the disturbed treatments to identify whether the disturbed communities separated from the AA

treatment in community composition. Strong separation of disturbed and undisturbed

communities would indicate novel community development in the disturbed treatments. The

PCA for the diatom communities captured 97.2% of community variability in the first two axes.

The first axis, responsible for 92.8% of variability, represented the tradeoff between

communities dominated by *Gomphonema* spp. and those dominated by *Nitzschia holsatica.* The

loadings for these two taxa on the first axis were -0.650 and 0.757, respectively. The second

eigenvector accounted for 4.4% of variability and corresponded to *Cocconeis* spp., colonial

*Fragilaria*, *Gomphonema* spp., and *Nitzschia holsatica* with loadings of -0.481, -0.428, 0.601, and 0.465, respectively.

The communities in the undisturbed treatment, AA, occupied a large area of the PCA space (Fig. 2A). Communities from the AA treatment spanned nearly the entire length of the first axis and had both the highest and lowest points on the second axis. This PCA indicates that communities within the AA treatment were highly variable, even in the context of the other, disturbed communities. Additionally, the majority of disturbed communities occur within the area spanned by the AA treatment, suggesting that there are no major differences in community composition between the AA treatment and the disturbed treatments.

*Bacterial Communities*

Bacterial community composition in the periphyton was determined using a PCR-based DNA fingerprinting method called Automated Ribosomal Intergenic Spacer Analysis (ARISA) (35). This method generates a measure of the relative abundance of each population that was amplified by PCR, allowing for rapid comparisons of many samples. Each detected amplicon corresponds to an operational taxonomic unit (OTU). We used a linear mixed model analysis with the same structure as the mixed model for the diatoms, but using the normalized ARISA-peak height as a measure of population relative abundance. Again, we used this analysis to compare the variability of taxa in communities experiencing disturbances to the variability of taxa in the undisturbed treatment, AA. The treatment effects and p values reported for the bacterial communities were obtained from this analysis.

The analyses for the bacterial communities used the residuals from detrended taxon CVs as the response variable. Detrending was performed because the CVs of the OTUs showed a

strong relationship with mean OTU abundances, whereby OTUs with higher mean abundances had lower CVs. To remove the effect of mean abundance, we fit an exponential model of the OTU CVs as a function of log(OTU mean abundances) (Fig. 1). We then used the residuals of this model as the metric of variability for each OTU, because OTUs with positive residuals were more variable than expected, whereas OTUs with negative residuals were less variable than expected. Because we used these model residuals instead of the OTU CVs, the effect sizes and standard errors are smaller in the bacterial analysis than in the diatom analysis.

For the bacterial communities, AA had the median level of variability out of the nine treatments. However, no treatment was significantly more variable than the AA treatment. No single disturbance at either Time 1 or Time 2 had significant effects on the variability of OTUs (Table 2). Thus, the OTUs in the treatments SA, AS, DA, and AD did not strongly differ in variability from those in the undisturbed treatment, AA. However, three out of the four terms for interactions between disturbances at Time 1 and disturbances at Time 2 were significant and negative; the treatments DD, SD, and SS had lower CVs than would have been predicted by the additive effects of disturbances at Time 1 and Time 2. Thus, the significant interaction terms show that the responses of the DD, SD, and SS treatments differed substantially from the independent effects of single disturbances. The effect sizes of these interactions show that DD, SD, and SS were the three least variable treatments (Table 2).

As with the diatom communities, we compared the bacterial community composition of the AA treatment to the disturbed treatments using a PCA. We evaluated whether the communities in the AA treatment separated from the communities in disturbed treatments to determine if there were consistent compositional differences between the undisturbed and disturbed communities. The PCA for the bacterial communities captured 51.2% of community

variability in the first two axes. The first axis accounted for 27.6% of community variability and the second axis accounted for 23.6% of the variability. The loadings on these two axes were primarily from the most abundant OTUs across all treatments.

The communities in the undisturbed treatment, AA, occurred in close proximity to disturbed communities in PCA space (Fig. 2B). Furthermore, similar to the diatom ordination, the AA treatment polygon overlapped with every other treatment polygon (Figures S6 and S7). Additionally, communities in the AA treatment fell along a wide range of the first axis, indicating that communities in this treatment showed substantial variability in community composition. Although many disturbed communities lay outside of the area encompassed by the undisturbed treatment, there was no strong separation between disturbed communities and the AA communities. As with the diatom communities, these results suggest that the disturbed communities did not consistently differ from the AA communities in terms of community composition.

*Comparing the Two Communities*

We used three different dissimilarity metrics (Sorensen, Euclidean, and Bray-Curtis) in Mantel tests to evaluate whether differences in the diatom communities were related to differences in the bacterial communities. There was no significant relationship between the bacterial and diatom communities for any of the three metrics used ($p = 0.540$, $p = 0.554$, and $p = 0.754$ for Sorensen, Euclidean, and Bray-Curtis, respectively).

We also compared the effects of the linear mixed models of the bacteria and diatom communities. We plotted the average treatment effects from the mixed models to compare how the same treatment affected the diatom communities and the bacterial communities that co-occurred on the slides (Fig. 3). We divided the plot into four quadrants by overlaying the grand

mean response of all nine treatments for the diatom and bacterial communities. Treatments with a response that was greater than the mean were relatively more variable, while treatments that fell below the mean response were relatively less variable. Treatments in quadrant *I* were above average in variability in both the diatom and the bacteria communities. Treatments in quadrant *II* were low in variability in the diatom communities, but high in variability in the bacteria communities. The reverse was true of the treatments that fell in quadrant *IV*. Finally, treatments in quadrant *III* were less variable than average for both the diatom and the bacteria communities.

The AA treatment was firmly inside quadrant *I*, the most variable quadrant. Conversely, the only two treatments to fall within the least variable quadrant were the doubly-disturbed treatments SD and DD. Furthermore, treatments experiencing the same disturbances, but in different orders, often appeared in different quadrants. The AD and DA disturbances experienced opposite effects, appearing in quadrants *II* and *IV*, respectively. Similarly, communities in the SD and DS treatments showed differing effects, particularly along the bacterial axis.

## Discussion

These experiments support the hypotheses that (i) disturbances decrease the variability of populations within diatom and bacteria communities and that (ii) multiple disturbances have interactive effects. For the diatom communities, every treatment that experienced a single disturbance had a significantly lower square root CV than the AA treatment. This consistent result shows that communities that were disturbed once became less variable than communities that were undisturbed. However, double disturbances did not necessarily cause the communities to become more consistent. Both the SS and the DD treatments had significant positive interactions, showing that these communities were more variable than would be expected based on the independent effects of the single disturbances. However, the treatments that experienced

different disturbances at the two time points (SD and DS) continued to become less variable as a result of the second disturbance. Thus, for the diatom communities, different disturbances continued to increase the consistency of the diatom communities, although the same recurring disturbance appeared to be saturating in effect. Therefore, the interactions between sequential disturbances were important to understanding community dynamics in treatments that were disturbed twice.

For the bacterial communities, no single disturbance had a significant effect on the variability of OTUs within the communities. However, three of the four treatments that experienced two disturbances had interactive effects, all of which led to lower population CVs. Thus, multiple disturbances to the bacterial communities generally created communities that were less variable than communities that were disturbed once. The strong interactions indicate that multiple disturbances had novel effects on the communities, such that the communities that experienced two disturbances demonstrated much different responses than the communities that only experienced one disturbance. Additionally, it appears that high levels of disturbance were necessary to generate changes in the bacterial communities, because the only treatments to show significant effects were disturbed twice. This was not surprising, as pelagic lake bacterial communities were previously found to be highly resilient to disturbances (25), and therefore may have recovered or experienced substantial turnover during the course of the experiment.

Despite differences in how the diatoms and the bacteria responded to individual treatments, there were several broad similarities between the diatom and bacterial responses to disturbances. For instance, none of the disturbed treatments in either the diatom or bacterial communities showed significantly greater variability than the undisturbed treatment, AA. Additionally, in both the diatom and the bacterial communities, at least two of the three least

variable treatments were highly disturbed, having experienced two disturbances (Fig. 3). These results suggest that the disturbances imposed on the periphyton communities acted as canalizing ecological drivers and constrained the variability of populations within the periphyton. However, there was no overall relationship between the response of the bacterial population variability and the response of the diatom population variability (Fig. 3). In several instances, treatments were more variable than average for either the diatom or bacterial community, but less variable than average in the other community. Thus, although disturbances generally decreased variability across all communities, there was no simple relationship between changes in the bacterial communities and the diatom communities on the same slides.

The differences in the diatoms and bacteria responses to individual disturbance treatments may be due to additional drivers of population variability within these two communities. For example, the strength and number of species interactions in a community can also be an important determinant of population variability (4). Although we have no estimates of species interactions in the diatom or bacterial communities, we note that the average strength and number of species interactions in these two communities may be different. This observation is based on the differing structure (richness, evenness) of the diatom and bacterial communities. Furthermore, we expect the rates of turnover to differ between the diatom and bacteria communities, with bacteria growing at a faster average rate. Thus, these varying growth rates could contribute to unequal rates of turnover between the two communities, which is an important factor mediating how quickly communities recover from disturbances (25). Understanding how these various drivers of population variability interact is necessary for predicting the variability of community processes, such as changes in biomass, production, or respiration (29).

Prior work has suggested that disturbances may mediate stochastic community assembly by enforcing a niche-based environmental filter (30). Our results from the diatom communities agree with this hypothesis, as the dominant taxa in disturbed communities have traits that may confer an advantage under the disturbed conditions. For instance, some *Gomphonema* species have been found to be tolerant of turbulent conditions, showing high abundances in water currents (31). Thus, they may have been particularly resistant to the water scouring disturbance. Similarly, *N. holsatica* is a small diatom that can become highly abundant in Icelandic lakes during the spring and summer (32). One hypothesis for the dominance of *N. holsatica* under the altered depth disturbance is that the species reproduced rapidly in the conditions of higher light due to its small size, and, therefore, high growth rate (33). Thus, we find support for the hypothesis that the harsh environmental conditions imposed by our experimental disturbances created an environmental filter, wherein taxa with functional traits favored by the disturbance could thrive under these conditions.

Although many microbial studies have demonstrated that community composition changes in response to environmental factors (e.g., 34-36), few have addressed the accuracy or repeatability of these results. However, studies that have evaluated the variability between disturbed microbial communities have found that there is often a high degree of similarity between strongly perturbed communities. Bell et al. (20) found that the bacterial communities following diesel contamination were similar in richness and composition following the disturbance. Similarly, Handley et al. (21) found that bacterial communities converged upon similar community compositions as a result of switching between acetate and lactate amendments. Therefore, these studies found that disturbances had repeatable effects on microbial systems, because disturbed communities were strikingly similar to one another. However, in

these two cases, the communities became more consistent partially as a result of novel communities developing under altered environmental conditions. In our study, populations became less variable in the absence of novel community development; in fact, the PCA polygons for the undisturbed bacterial and the diatom communities overlap substantially with every other treatment. In this case, disturbances increased the consistency of microbial communities by placing tighter constraints on community composition.

Many studies in microbial ecology have sought to quantify the degree to which communities are shaped by stochastic versus deterministic processes (37, 38). The main deterministic process discussed is varying selection strength on microbial taxa, usually as the result of environmental or biotic stress (37, 39). Selection is named as a deterministic force under the assumption that consistent and differential selection will eventually lead to the same final community composition (40). Conversely, colonization and drift are two stochastic forces that are important in community assembly (9, 41). We find evidence in our experiments for the stochastic effects of colonization by observing the wide variability of populations in AA communities, which was presumably determined by the stochasticity of colonizers on the Plexiglas slides. Additionally, drift may be a particularly important force in communities if there is a high degree of functional redundancy (9), which can lead to communities that vary in the abundances of ecologically equivalent taxa (42). Bacterial communities, in particular, have been hypothesized to have relatively high functional redundancy of taxa due to their high species richness (43). Thus, if bacterial communities are predisposed to experience greater compositional drift due to the existence of ecologically equivalent taxa, then bacterial populations should be expected to be more variable than populations within communities with fewer ecological equivalents. This offers another explanation as to why the bacterial populations in our

experiment showed no significant decreased in variability after single experimental disturbances, whereas the diatom populations did show significant decreases in variability.

Our study suggests that environmental stressors can indeed act as a deterministic force in microbial communities, because the communities stressed by our disturbances became less variable after experiencing the disturbances. These results were consistent across the two experimental disturbances imposed, despite the different nature and time scale of the two disturbances. Specifically, the water scouring disturbance was a perturbation of high impact over a short period of time (pulse disturbance), whereas the altered depth disturbance was a sustained perturbation (press disturbance). The similar response of the communities to both these disturbances is in line with a recent review showing that a high proportion of microbial communities are sensitive to both press and pulse disturbances; of the experiments reviewed, 92 of 112 microbial communities showed a change in composition or function in response to a pulse disturbance, and 141 of 178 communities changed in response to a press disturbance (25). Thus, the variability of microbial communities may be a useful indicator of the degree to which the communities are influenced by stochastic or deterministic processes, because many microbial communities are sensitive to disturbance. However, studying the variability of populations requires a high degree of replication, which is often lacking in microbial ecology (44). Thus, characterizing the magnitude of microbial community variability, and of the forces contributing to this variability, requires an amount of replication that is seldom found in microbial studies.

In addition to the experimental disturbances, there are many other possible factors that could have influenced the variability of populations within the periphyton. For instance, the periphyton communities experienced environmental variability throughout the duration of the experiment due to natural weather conditions and small-scale variability in environmental forces.

Acknowledging this environmental stochasticity, we intentionally implemented experimental disturbances that were more extreme than the natural variability we observed during this time period. Additionally, we did not account for colonization of diatoms or bacteria after disturbances were implemented, which may have generated additional variability in these communities. However, because the Plexiglas slides were re-randomized between disturbances, we expect systematic bias from immigration to be minimal between treatments. Furthermore, the diatom and bacterial datasets are complimentary in their strengths; the diatom data were obtained through direct counts, meaning that there is high accuracy in identification, although only a subsample of the community was measured. Conversely, nearly the entire bacterial community was sampled, but with some degree of bias from using ARISA (45). Thus, because the two datasets were obtained using different methods, we are confident that the similarities in the results are not an artifact of our methodology.

Prediction of microbial communities is an oft-cited goal of microbial ecology. However, predictive models can only be accurate if the process they are describing is inherently repeatable. For instance, statistical models will only produce a good fit to microbial community composition data if these microbial communities show consistent responses to environmental drivers. The results of these experiments indicate that microbial communities do show repeatability in their response to environmental stress, because communities became more similar to one another after experiencing the same disturbance. This finding could be tested in other systems by examining whether predictive models of bacterial community composition (e.g. 14) have lower error when modeling disturbed communities. These results suggest that changes to microbial communities could be modeled using abiotic drivers as predictors. However, the diatom and bacterial communities varied in susceptibility to environmental forcing, as the effects of the same

treatment on the two communities often differed. Thus, the abiotic drivers that are the best predictors of community composition are likely to vary across different ecological communities and ecosystems, as might be expected from first principles.

**Materials and Methods**

*Experimental Manipulations*

These experiments were performed in Lake Myvatn, a shallow, eutrophic (external loading of $1.4 \text{ g} \cdot \text{m}^{-2} \cdot \text{y}^{-1}$ of N and $1.5 \text{ g} \cdot \text{m}^{-2} \cdot \text{y}^{-1}$ of P; net algal production of $222 \text{ g} \cdot \text{m}^{-2} \cdot \text{y}^{-1}$ of C, [46]) lake in northeast Iceland (65° 40' N, 17° 00' W, [28]). We allowed periphyton to colonize the Plexiglas substrate for 20 days before beginning the disturbance manipulations. During this period of colonization, 108 Plexiglas slides (6cm x 8 cm) were suspended in Lake Myvatn at 3m depth, which was approximately 0.3m from the sediment surface.

Disturbances were implemented at two time points. At the first time point, day 20, we randomly assigned 36 of the 108 slides to each of the following conditions: ambient (no disturbance), altered depth disturbance (relocation to 0.5m depth), or water scouring disturbance. Weekly water column profiles showed that Secchi depth during the summer of 2013 varied between 1.5m and 3.3m, with 11-33% of surface photosynthetically active radiation (PAR) reaching 3m depth. We chose these two disturbances (altered depth and water scouring) because they mimic natural disturbances to periphyton communities in the lake due to the high wind events that are common at our study site (28, 47). During these high wind events, periphyton communities may experience a change in depth due to resuspension in the water column, or individuals might be scoured from biofilms due to fast water currents.

Similarly, on day 25, we again randomized the slides into three groups and manipulated the slides with the disturbances described above, incubating the slides for another 5 days. On day

30 of the experiment, we retrieved the Plexiglas slides from the lake and froze the slides at -20C
until further processing. Additional details about the experimental manipulations are provided in
the Supplementary Materials (Figure S1).

*Community composition analysis*

Slides were removed briefly from the freezer to obtain diatom counts on a microscope
before being frozen again. Diatoms were identified to the lowest taxonomic resolution possible,
which was genus or species. A minimum of 500 individuals per slide were identified by counting
half transects across slides. The mean number of individuals identified per sample was 1 063, for
a total of 114 843 individuals across the 108 Plexiglas slides. We then transported the slides to
Madison, Wisconsin, USA, for analysis of the bacterial communities in the periphyton using
*Automated Ribosomal Intergenic Spacer Analysis (ARISA)* (48, 49). Briefly, DNA was extracted
from periphyton biomass that was scraped from the slides and this DNA was used as template for
PCR to amplify the intergenic region between the 16S and 23S rRNA genes in bacteria.
Amplicons were separated by capillary electrophoresis and used to define operational taxonomic
units (OTUs). Additional details about community composition analysis are provided in the
Supplementary Materials.

**Statistical Methods**

*Diatom Communities*

Because some taxa were rare, and therefore were inconsistently present in samples, we
analyzed only the 8 most common diatom taxa (*Nitschia holsatica*, *Cymbella spp.*, *Synedra spp.*,
*Gomphonema spp.*, *Rhoicosphenia spp.*, *Cocconeis spp.*, colonial *Fragilaria spp.* and singular

*Fragilaria spp.)*. Together, these 8 taxa accounted for 99.4% of all individuals counted. We standardized all data to densities of each taxon per cm$^2$. For each of the nine treatments, we calculated the coefficient of variation (CV) for each of the 8 taxa across the 12 slides in that treatment. We chose the CV as the indicator of population variability because it did not change in response to the mean abundance of the taxa and because it showed homogeneity in variance between treatments. Additionally, the CV integrates across all 12 slides within a treatment and mitigates the effects of any single anomalous communities. We transformed the CVs by taking their square root because the distribution of CVs was slightly skewed toward larger values.

We analyzed square root CVs using a linear mixed effects model. The four predictor variables ( $X_{D1}$ , $X_{S1}$ , $X_{D2}$ , and $X_{S2}$ ) were binary vectors corresponding to whether or not the taxon was in a treatment that received the depth disturbance at Time 1, the scouring disturbance at Time 1, the depth disturbance at Time 2, or the scouring disturbance at Time 2 (Eq. 1). We also included all interactions between these four predictor variables to assess the interactive effects of multiple disturbances. We recognized that the taxa may have different inherent levels of population variability, and so we included a random intercept by taxon, denoted by $\alpha_{taxon}$ . This term assumes that the square root CVs of taxa are normally distributed, but only estimates the distribution from which the square root CVs are drawn, rather than an effect for each taxon.

$$
\begin{aligned}
\sqrt{CV} &= \beta_0 + \beta_1 X_{D1} + \beta_2 X_{S1} + \beta_3 X_{D2} + \beta_4 X_{S2} \\
&+ \beta_5 (X_{D1} \bullet X_{D2}) + \beta_6 (X_{D1} \bullet X_{S2}) + \beta_7 (X_{S1} \bullet X_{D2}) + \beta_8 (X_{S1} \bullet X_{S2}) \\
&+ \alpha_{taxon} + \varepsilon \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad \text{Eq. 1}
\end{aligned}
$$

$$\alpha_{taxon} \sim N(0, \sigma_{intercept}^2)$$

$$\varepsilon \sim N(0, \sigma_{error}^2)$$

*Bacterial Communities*

We removed two bacterial samples from our analyses due to their anomalously low

diversities, resulting in 106 bacterial samples. When evaluating population variability, we

analyzed OTUs that were present in at least 30 samples, which included 55 OTUs. For each of

the 9 treatments, we calculated the CV for each OTU within that treatment. However, the CVs of

the OTUs were correlated with mean OTU abundance, with highly abundant OTUs generally

having lower CVs. This is a common pattern when data are relativized, due to the

heteroskedasticity of binomial data (50). To account for this expected pattern, we detrended the

data by fitting the CVs as a negative exponential function of the log(OTU mean relative

abundance) (Fig. 1). We then used the residuals of this function as the response variable in our

analyses. Points above the fitted relationship (positive residuals) are OTUs that were more

variable than would be expected, after the effect of abundance was removed, whereas points

below the line (negative residuals) were less variable than would be expected. We then analyzed

the residuals using an analogous statistical model as was used with the diatom data (Eq. 2).

$$
\begin{aligned}
Residuals = \; & \beta_0 + \beta_1 X_{D1} + \beta_2 X_{S1} + \beta_3 X_{D2} + \beta_4 X_{S2} \\
& + \beta_5 (X_{D1} \bullet X_{D2}) + \beta_6 (X_{D1} \bullet X_{S2}) + \beta_7 (X_{S1} \bullet X_{D2}) + \beta_8 (X_{S1} \bullet X_{S2}) \\
& + \alpha_{OTU} + \varepsilon \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{Eq. 2}
\end{aligned}
$$

$$\alpha_{OTU} \sim N(0, \sigma^2_{intercept})$$

$$\varepsilon \sim N(0, \sigma^2_{error})$$

Prior to fitting the statistical model, we removed the 6 outliers that were greater than 3

standard deviations from the mean of the residuals. However, the model was robust to these

outliers and identified the same treatments as significant when the outliers were included. The

results were also robust to changes in the frequency cutoff used to determine the number of

OTUs included; the model identified the same treatments as significant when varying the cutoff

for inclusion in the analysis between presence in at least 20 samples and presence in at least 40 samples. Additional information on statistical methods and diagnostics can be found in the Supplementary Online Materials (Figs. S2 – S5).

*Principal Components Analysis of Communities*

We performed Principal Components Analyses (PCAs) on the diatom and the bacterial communities. Our main goal for these analyses was to evaluate whether the composition of the disturbed communities consistently differed from the composition of undisturbed communities. Because there was a wide range in the mean densities of diatoms on the slides, we transformed the diatom counts into relative abundances before running this analysis. Again, we used only the 8 most common taxa in the diatom PCA. Similarly, in the PCA of the bacterial communities, we removed all OTUs present in fewer than 30 samples.

*Comparing the Two Communities*

To assess whether there were correlations between the diatom communities and the bacterial communities, we performed Mantel tests on the 106 slides for which we had data on both the diatom and bacterial communities. We used these tests to determine if changes to either the diatom or the bacterial community on a slide could predict changes in the other community.

**References**

1. **Colwell RK**. 1974. Predictability, Constancy, and Contingency of Periodic Phenomena. Ecology **55**:1148–1153.

2. **McGrady-Steed J**, **Harris PM**, **Morin PJ**. 1997. Biodiversity regulates ecosystem predictability. Nature **390**:162–165.

3. **Benedetti-Cecchi L**. 2003. The importance of the variance around the mean effect size of ecological processes. Ecology **84**:2335–2346.

4. **Kilpatrick AM**, **Ives AR**. 2003. Species interactions can explain Taylor's power law for ecological time series. Nature **422**:65–68.

5. **Turchin P.** 1995. Population Regulation: Old Arguments and a New Synthesis. p 19-39. *In* Cappuccino N, Price PW (ed), Population Dynamics: New Approaches and Synthesis. Academic Press.

6. **Fraterrigo JM**, **Rusak JA**. 2008. Disturbance-driven changes in the variability of ecological patterns and processes. Ecology Letters **11**:756–770.

7. **Forrest J**, **Arnott SE**. 2007. Variability and predictability in a zooplankton community: The roles of disturbance and dispersal. Écoscience **14**:137–145.

8. **Turner MG**, **Romme WH**, **Gardner RH**, **O'Neill RV**, **Kratz TK**. 1993. A revised concept of landscape equilibrium: Disturbance and stability on scaled landscapes. Landscape Ecol **8**:213–227.

9. **Nemergut DR**, **Schmidt SK**, **Fukami T**, **O'Neill SP**, **Bilinski TM**, **Stanish LF**, **Knelman**

**JE**, **Darcy JL**, **Lynch RC**, **Wickey P**, **Ferrenberg S**. 2013. Patterns and Processes of Microbial Community Assembly. Microbiol Mol Biol Rev **77**:342–356.

10. **Ranta E**, **Kaitala V**, **Lindström J**, **Helle E**. 1997. The Moran Effect and Synchrony in Population Dynamics. Oikos **78**:136–142.

11. **Tilman D**. 1999. The ecological consequences of changes in biodiversity: a search for general principles. Ecology **80**:1455–1474.

12. **Fernandez AS**, **Hashsham SA**, **Dollhopf SL**, **Raskin L**, **Glagoleva O**, **Dazzo FB**, **Hickey RF**, **Criddle CS**, **Tiedje JM**. 2000. Flexible Community Structure Correlates with Stable Community Function in Methanogenic Bioreactor Communities Perturbed by Glucose. Appl Environ Microbiol **66**:4058–4067.

13. **Roelke D**, **Augustine S**, **Buyukates Y**. 2003. Fundamental Predictability in Multispecies Competition: The Influence of Large Disturbance. The American Naturalist **162**:615–623.

14. **Larsen P**, **Hamada Y**, **Gilbert J**. 2012. Modeling microbial communities: Current, developing, and future technologies for predicting microbial community interaction. Journal of Biotechnology **160**:17–24.

15. **Rykiel EJ**. 1985. Towards a definition of ecological disturbance. Australian Journal of Ecology **10**:361–365.

16. **Bissett A**, **Brown MV**, **Siciliano SD**, **Thrall PH**, **Holyoak M**. 2013. Microbial community responses to anthropogenically induced environmental change: towards a systems approach. Ecology Letters **16**:128–139.

17. **de Vries FT**, **Shade A**. 2013. Controls on soil microbial community stability under climate change. Front Microbiol **4**.

18. **Crowther TW**, **Maynard DS**, **Leff JW**, **Oldfield EE**, **McCulley RL**, **Fierer N**, **Bradford MA**. 2014. Predicting the responsiveness of soil biodiversity to deforestation: a cross-biome study. Glob Change Biol **20**:2983–2994.

19. **Holden SR**, **Treseder KK**. 2013. A meta-analysis of soil microbial biomass responses to forest disturbances. Front Microbiol **4**.

20. **Bell TH**, **Yergeau E**, **Maynard C**, **Juck D**, **Whyte LG**, **Greer CW**. 2013. Predictable bacterial composition and hydrocarbon degradation in Arctic soils following diesel and nutrient disturbance. ISME J **7**:1200–1210.

21. **Handley KM**, **Wrighton KC**, **Miller CS**, **Wilkins MJ**, **Kantor RS**, **Thomas BC**, **Williams KH**, **Gilbert JA**, **Long PE**, **Banfield JF**. 2014. Disturbed subsurface microbial communities follow equivalent trajectories despite different structural starting points. Environ Microbiol.

22. **Turner MG**, **Collins SL**, **Lugo AL**, **Magnuson JJ**, **Rupp TS**, **Swanson FJ**. 2003. Disturbance Dynamics and Ecological Response: The Contribution of Long-Term Ecological Research. BioScience **53**:46–56.

23. **White PS**, **Jentsch A**. 2001. The Search for Generality in Studies of Disturbance and Ecosystem Dynamics, p. 399–450. *In* Esser, K, Lüttge, U, Kadereit, JW, Beyschlag, W (eds.), Progress in Botany. Springer Berlin Heidelberg.

24. **Kerkhoff AJ**, **Enquist BJ**. 2007. The Implications of Scaling Approaches for Understanding Resilience and Reorganization in Ecosystems. BioScience **57**:489–499.

25. **Shade A**, **Peter H**, **Allison SD**, **Baho DL**, **Berga M**, **Burgmann H**, **Huber DH**, **Langenheder S**, **Lennon JT**, **Martiny JBH**, **Matulich KL**, **Schmidt TM**, **Handelsman J**. 2012. Fundamentals of Microbial Community Resistance and Resilience. Front Microbiol **3**.

26. **Harley CDG**, **Paine RT**. 2009. Contingencies and compounded rare perturbations dictate sudden distributional shifts during periods of gradual climate change. PNAS **106**:11172–11176.

27. **Paine RT**, **Tegner MJ**, **Johnson EA**. 1998. Compounded Perturbations Yield Ecological Surprises. Ecosystems **1**:535–545.

28. **Einarsson Á**, **Stefánsdóttir G**, **Jóhannesson H**, **Ólafsson JS**, **Gíslason GM**, **Wakana I**, **Gudbergsson G**, **Gardarsson A**. 2004. The ecology of Lake Myvatn and the River Laxá: Variation in space and time. Aquatic Ecology **38**:317–348.

29. **Doak DF**, **Bigger D**, **Harding EK**, **Marvier MA**, **O'Malley RE**, **Thomson D**. 1998. The Statistical Inevitability of Stability‐Diversity Relationships in Community Ecology. The American Naturalist **151**:264–276.

30. **Chase JM**. 2007. Drought mediates the importance of stochastic community assembly. PNAS **104**:17430–17434.

31. **Whitford LA**. 1960. The Current Effect and Growth of Fresh-Water Algae. Transactions of

the American Microscopical Society **79**:302–309.

32. **Jónasson PM**, **Adalsteinsson H**, **St. Jónsson G**. 1992. Production and Nutrient Supply of Phytoplankton in Subarctic, Dimictic Thingvallavatn, Iceland. Oikos **64**:162–187.

33. **Chisholm SW**. 1992. Phytoplankton Size, p. 213–237. *In* Falkowski, PG, Woodhead, AD, Vivirito, K (eds.), Primary Productivity and Biogeochemical Cycles in the Sea. Springer US.

34. **Prober SM**, **Leff JW**, **Bates ST**, **Borer ET**, **Firn J**, **Harpole WS**, **Lind EM**, **Seabloom EW**, **Adler PB**, **Bakker JD**, **Cleland EE**, **DeCrappeo NM**, **DeLorenze E**, **Hagenah N**, **Hautier Y**, **Hofmockel KS**, **Kirkman KP**, **Knops JMH**, **La Pierre KJ**, **MacDougall AS**, **McCulley RL**, **Mitchell CE**, **Risch AC**, **Schuetz M**, **Stevens CJ**, **Williams RJ**, **Fierer N**. 2015. Plant diversity predicts beta but not alpha diversity of soil microbes across grasslands worldwide. Ecol Lett **18**:85–95.

35. **Fierer N**, **Jackson RB**. 2006. The diversity and biogeography of soil bacterial communities. PNAS **103**:626–631.

36. **Lauber CL**, **Hamady M**, **Knight R**, **Fierer N**. 2009. Pyrosequencing-Based Assessment of Soil pH as a Predictor of Soil Bacterial Community Structure at the Continental Scale. Appl Environ Microbiol **75**:5111–5120.

37. **Dini-Andreote F**, **Stegen JC**, **Elsas JD van**, **Salles JF**. 2015. Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. PNAS 201414261.

38. **Zhou J**, **Deng Y**, **Zhang P**, **Xue K**, **Liang Y**, **Nostrand JDV**, **Yang Y**, **He Z**, **Wu L**, **Stahl DA**, **Hazen TC**, **Tiedje JM**, **Arkin AP**. 2014. Stochasticity, succession, and environmental perturbations in a fluidic ecosystem. Proceedings of the National Academy of Sciences **111**:E836–E845.

39. **Valverde A**, **Makhalanyane TP**, **Cowan DA**. 2014. Contrasting assembly processes in a bacterial metacommunity along a desiccation gradient. Front Microbiol **5**.

40. **Vellend M**. 2010. Conceptual Synthesis in Community Ecology. The Quarterly Review of Biology **85**:183–206.

41. **Vanschoenwinkel B**, **Waterkeyn A**, **Jocqué M**, **Boven L**, **Seaman M**, **Brendonck L**. 2010. Species sorting in space and time—the impact of disturbance regime on community assembly in a temporary pool metacommunity. Journal of the North American Benthological Society **29**:1267–1278.

42. **Leibold MA**, **McPeek MA**. 2006. Coexistence of the Niche and Neutral Perspectives in Community Ecology. Ecology **87**:1399–1410.

43. **Bell T**, **Newman JA**, **Silverman BW**, **Turner SL**, **Lilley AK**. 2005. The contribution of species richness and composition to bacterial services. Nature **436**:1157–1160.

44. **Prosser JI**. 2010. Replicate or lie. Environmental Microbiology **12**:1806–1810.

45. **van Dorst J**, **Bissett A**, **Palmer AS**, **Brown M**, **Snape I**, **Stark JS**, **Raymond B**, **McKinlay J**, **Ji M**, **Winsley T**, **Ferrari BC**. 2014. Community fingerprinting in a sequencing world. FEMS Microbiol Ecol.

46. **Ólafsson J**. 1979. The Chemistry of Lake Mývatn and River Laxá. Oikos **32**:82–112.

47. **Ólafsson J**. 1979. Physical Characteristics of Lake Mývatn and River Laxá. Oikos **32**:38–66.

48. **Fisher MM**, **Triplett EW**. 1999. Automated Approach for Ribosomal Intergenic Spacer Analysis of Microbial Diversity and Its Application to Freshwater Bacterial Communities. Appl Environ Microbiol **65**:4630–4636.

49. **Jones SE**, **Shade AL**, **McMahon KD**, **Kent AD**. 2007. Comparison of Primer Sets for Use in Automated Ribosomal Intergenic Spacer Analysis of Aquatic Bacterial Communities: an Ecological Perspective. Appl Environ Microbiol **73**:659–662.

50. **Pampel FC**. 2000. Logistic Regression: A Primer. SAGE.

**Table 1: Mixed Model Results for Diatoms**

Results of the linear mixed model using disturbances at Time 1 and Time 2 as predictors of the taxon-level variability (as given by the square root of the taxon CVs) of diatom communities from the nine experimental treatments. Disturbance effect estimates are given in comparison to the undisturbed treatment, AA, which is why there is no p value estimate for the AA treatment. Each of the single disturbances at Time 1 and Time 2 significantly reduced the average taxon square root CV. There were significant positive interactions for communities that received the same disturbance at Time 1 and Time 2, corresponding to the DD and SS treatments. No p value was calculated for the random effect, because we were not interested in testing how much variability was explained by differences between taxa.

| Disturbance | Estimated Effect | p value |
|---|---|---|
| Intercept (AA) | 1.251 | n/a |
| Time 1: D | -0.156 | 0.0330  * |
| Time 1: S | -0.155 | 0.0343  * |
| Time 2: D | -0.230 | 0.0016  ** |
| Time 2: S | -0.158 | 0.0307  * |
| Time 1: D * Time 2: D | 0.207 | 0.0458  * |
| Time 1: S * Time 2: D | 0.102 | 0.3220 |
| Time 1: D * Time 2: S | 0.109 | 0.2899 |
| Time 1: S * Time 2: S | 0.217 | 0.0363  * |
| **Random Effect** | **Estimated Effect** | **p value** |
| Taxon | 0.0240 | n/a |

**Table 2: Mixed Model Results for Bacteria**

Results of the linear mixed model using disturbances at Time 1 and Time 2 as predictors of the OTU-level variability (as given by the residuals of OTU CVs) of bacterial communities from the nine experimental treatments. As in Table 1, disturbance effect estimates are given in comparison to the undisturbed treatment, AA. There were no significant effects of single disturbances on the variability of OTUs at Time 1 or Time 2. However, there were significant negative interactions between three doubly-disturbed treatments, such that treatments DD, SD, and SS were less variable than would have been expected. No p value was calculated for the random effect, because we were not interested in testing how much variability was explained by differences between taxa.

| Disturbance | Estimated Effect | p value |
|---|---|---|
| Intercept (AA) | 0.0163 | n/a |
| Time 1: D | -0.0450 | 0.3262 |
| Time 1: S | 0.0565 | 0.2159 |
| Time 2: D | 0.0276 | 0.1510 |
| Time 2: S | 0.0116 | 0.5472 |
| Time 1: D * Time 2: D | -0.143 | 0.0282 * |
| Time 1: S * Time 2: D | -0.228 | <0.001 *** |
| Time 1: D * Time 2: S | 0.022 | 0.7350 |
| Time 1: S * Time 2: S | -0.176 | 0.0065 ** |
| **Random Effect** | **Estimated Effect** | **p value** |
| OTU | 0.0459 | n/a |

**Figure 1:**
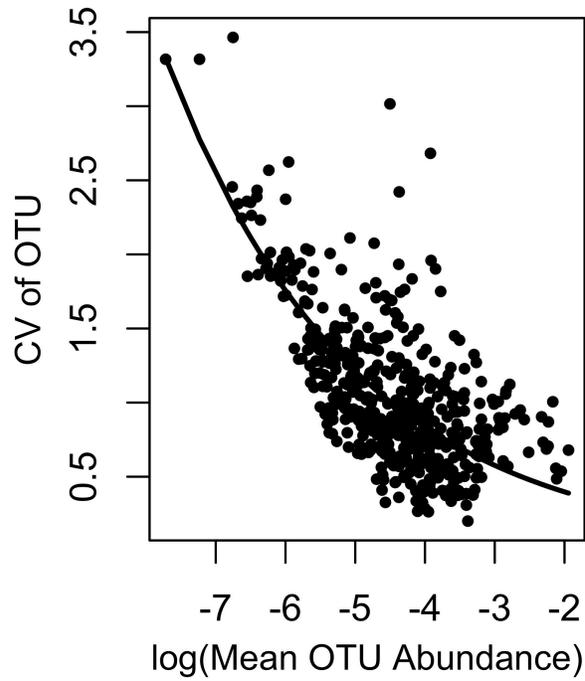


**Figure 1:** We detrended the CVs of OTUs from the ARISA data because the CVs were strongly related to mean OTU relative abundance. We expected the CVs of the OTUs to decrease as OTUs became more abundant. Thus, we fit an exponential function to the data and used the residuals of this relationship in the subsequent mixed model.

**Figure 2:**
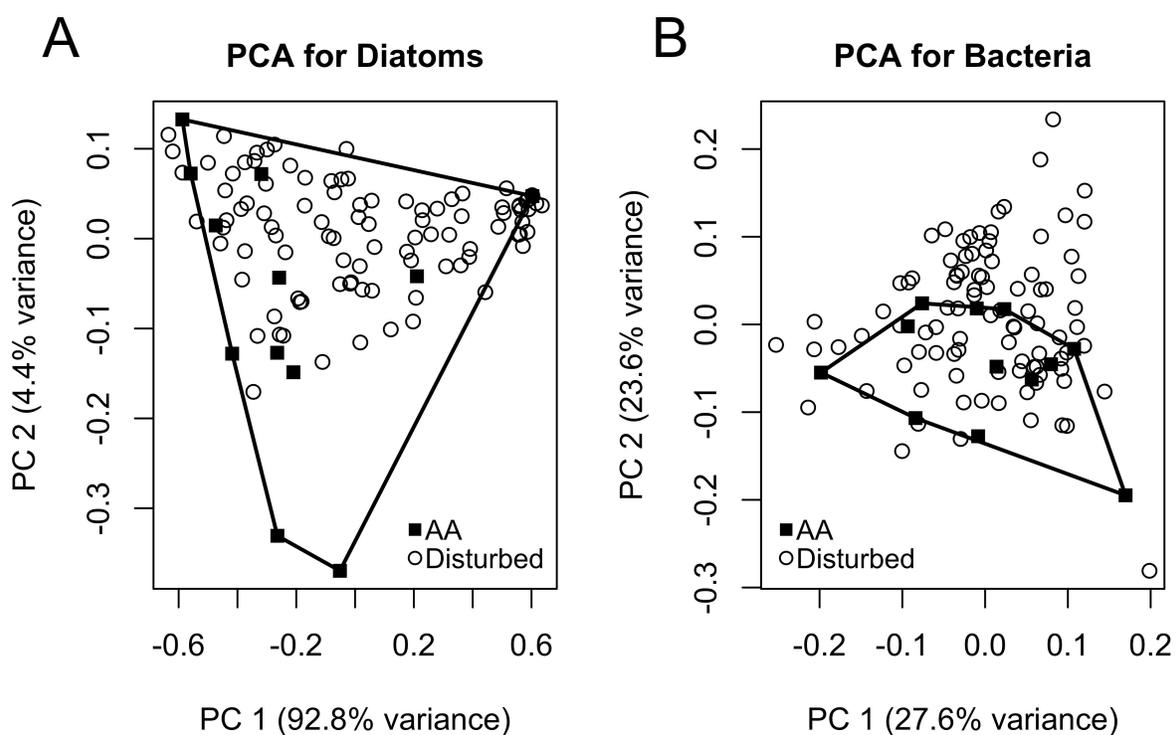


**Figure 2:** A) Principal components analysis of the diatom communities showed that the undisturbed treatment, AA, spanned most of the space occupied by the communities in the nine treatments. The majority of disturbed communities fell within the bounds of the AA communities, showing a lack of separation between the AA treatment and the disturbed treatments. The first and second axes together account for 97.2% of community variation. The polygon depicted shows the convex hull of the AA points, which is constructed by drawing the minimum number of connections between points to encapsulate the entire set of AA points. B) Results from the principal components analysis of the bacterial communities show that there is no strong differentiation between the community composition of the undisturbed treatment, AA, and that of the disturbed treatments. Additionally, the AA treatment covers a wide range of the PC 1 axis, which is the axis that explains the most variability between bacterial communities.

The first and second axes together account for 51.2% of community variation. As above, the polygon depicted shows the convex hull of the AA points.

**Figure 3:**
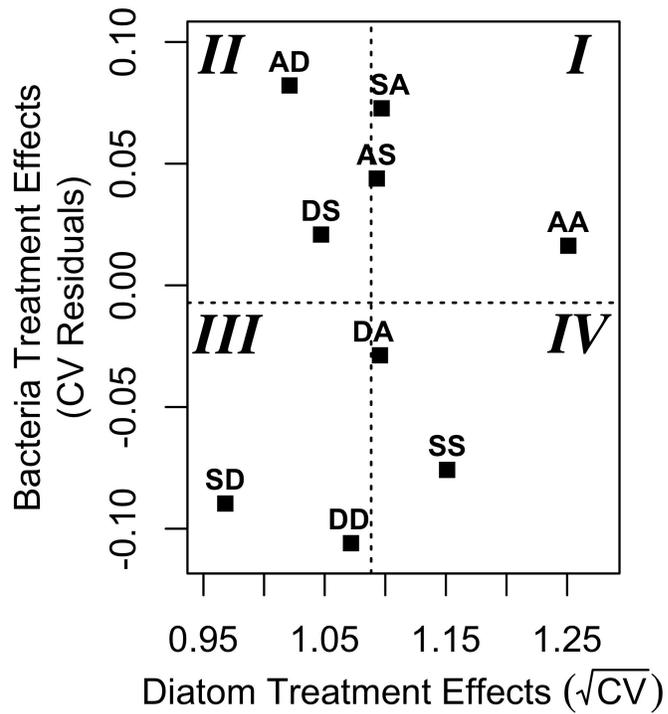


**Figure 3:** This plot shows the average variability of each treatment in the diatom and bacteria communities, as obtained from the mixed models. The dashed lines show the overall mean responses for the diatom and the bacteria treatments, such that treatments with a value higher than the mean are comparatively more variable. The AA treatment falls in the most variable portion of the plot (quadrant *I*), whereas the two communities that were least variable (quadrant *III*) were disturbed twice.

Supplementary Online Material

**Title:** Environmental disturbances increase the predictability of microbial communities within periphyton

**Authors:** Cristina M. Herren[1], Kyle C. Webert[2], and Katherine D. McMahon[3]

**Author affiliations:**

[1]Freshwater and Marine Sciences Program, University of Wisconsin - Madison, Madison, Wisconsin, USA

cherren@wisc.edu

[2]Department of Zoology, University of Wisconsin - Madison, Madison, Wisconsin, USA

[3]Departments of Bacteriology and Civil and Environmental Engineering, University of Wisconsin - Madison, Madison, Wisconsin, USA

**Additional Materials and Methods**

*Experimental Manipulations*

Periphyton was chosen as a study system because of the high species diversity of benthic algae in periphyton and the ability to analyze the co-occurring algal and bacterial communities. Additionally, diatoms at Lake Myvatn are particularly important to the lake food web. Lake Myvatn is a shallow, eutrophic lake that has high diatom production due to elevated concentrations of nutrients in the lake's groundwater inputs (Einarsson et al 2004).

Plexiglas slides (6cm x 8cm) were used as the substrate for periphyton growth. Two holes were drilled in the Plexiglas in order to attach slides to metal racks. Twelve replicates of each of the 9 treatments (108 total slides) were uniquely labeled and distributed randomly among six metal racks (18 slides per rack). Slides were randomized before the start of the experiment and before each of the two disturbances (Fig. S1). The bottom of the rack was attached to an anchor that sank into the lake sediment, and a line with a buoy was attached to the top of the rack to suspend the rack vertically. On this top line, there was also a smaller stabilization buoy secured 0.5m above the rack. All racks were deployed within a 5m radius.

We took care to minimally perturb the periphyton communities aside from the two disturbance treatments described. During disturbance manipulations, we retrieved all 6 racks and placed the slides on new racks corresponding to the treatment they were designated to receive at the next time point. In the time between removal from the initial rack and attachment to the next rack, each slide was stored individually in a covered container filled with lake water. The 36 slides receiving no disturbance were distributed across 2 racks and were simply replaced back into the water column at 3m depth. The 36 slides receiving the water scouring disturbance were distributed across 2 racks that were dragged through the water at 20-25cm/s for 10 minutes,

simulating a strong current in the lake due to a high wind event. After this disturbance, these racks were also replaced into the lake at a depth of 3m. The last set of 36 slides, which received the altered depth disturbance, were distributed across 2 racks and were replaced in the water column at a depth of 0.5m. These two racks remained at this shallower depth for 5 days, until the time of the second disturbance.

*Diatom Counts*

Slide counts were performed with a Leica compound microscope at 400x magnification. The vast majority (>99%) of algae on the slides were diatoms. In order to account for potential effects of spatial heterogeneity within each slide, half transects were counted across the Plexiglas slide, and only the center 6cm x 6cm area was counted. Diatoms were identified to the lowest taxonomic resolution possible, which was genus or species. Half transects were counted completely until a minimum of 500 individuals were identified. The mean number of individuals counted per sample was 1063, with 114,843 total individuals counted in the 108 samples. Seventeen taxonomic groupings were differentiated, and fewer than 1% of cells were not identifiable. Slides were again frozen after counting.

*Bacteria Analysis Using ARISA*

Slides were transported back to Madison, Wisconsin, USA, for analysis of the bacterial communities in the periphyton using *Automated Ribosomal Intergenic Spacer Analysis (ARISA)*. The same 6cm x 6cm area of the slide that was counted for periphyton was scraped with a sterile razor blade and transferred to a microcentrifuge tube. DNA from these samples was extracted using a xanthogenate-phenol-chloroform protocol described elsewhere by Miller and McMahon

(2011). We then used PCR with universal bacterial primers 1406f (5'-
TGYACACACCGCCCGT-3') and 23Sr (5'-GGGTTBCCCCATTCRG-3') to amplify the
intergenic spacer region between 16S and 23S of the bacteria in these samples. PCR reactions
used 5uL of 10x buffer, 2uL MgCl2, 1.25uL of dNTPs, 1uL of each primer, 1uL of template
DNA, 0.25uL of DNA polymerase, and 13.5uL of water. Samples were analyzed using
denaturing capillary electrophoresis using an ABI 3730 at the University of Wisconsin
Biotechnology Center. ARISA output was calibrated against a 100-2000bp standard
(Bioventures) and was analyzed using the GeneMarker v 1.5 software (SoftGenetics LLC) and
custom R scripts (Jones and McMahon 2009). The output of this software resulted in relative
abundance tables of the operational taxonomic units (OTUs) present in the samples.

**Additional Statistical Methods and Diagnostics**

*Diagnostics for Mixed Models*

In order to validate that the mixed model approach was appropriate for these data, we ran several diagnostic tests on the diatom data. First, we ensured that the measure of variability (the taxon CV) was not biased by mean taxon abundance. Using a linear regression, we found no effect of mean abundance on the CV of these populations (t = .34, p = 0.74, Fig. S2).

Similarly, we plotted the random effects estimated in the mixed model to identify whether these fitted effects were biased by the mean abundances of the taxa. We found no relationship between the estimated random effects and the average log mean abundance of the diatoms (Fig. S3). Thus, rare taxa were not more variable than common taxa, suggesting that sample sizes of the taxa were sufficiently large that the variability of taxa was not substantially influenced by sampling error.

Additionally, we checked that the residuals from the mixed model were not biased by treatment and were approximately normally distributed (Fig. S4).

*Additional Statistical Methods and Validation*

In order to assess whether the results of our analyses were robust to statistical methodology, we validated our results using a secondary analysis. Here, we present the results of this supplementary analysis conducted on the diatom data. This test gave similar results to the mixed model shown in the main text, as it also showed that the disturbed treatments were significantly lower in variability than the control treatment, AA.

In this second analysis, we bootstrap a mixed model similar to the one presented in the main text in order to obtain empirical p values for treatment effects, rather than relying upon p values obtained directly from the mixed model. We used this approach to verify that the significant treatment effects were not the result of a violation of the assumptions of a mixed model. Specifically, we were concerned that using taxon CVs as the statistical unit may inflate the probability of finding spurious statistical significance; the taxa were all part of the same communities on the Plexiglas slides, which could lead to interdependence among taxon CVs. Thus, to address this concern, we used a bootstrapping approach to account for any effect the co-occurrence of taxa on the same slide might have on the estimated treatment effects. We used this null model obtained from the bootstrap to discern significance, rather than relying upon the theoretical distribution of treatment effects.

To validate our statistical methods, we compared the p values for treatment effects that were obtained from the two methodologies described above. For ease of comparison, the model that we used for this validation was the simplest possible mixed model for our diatom data; we analyzed the square root CVs of the diatom taxa as a function of the 9 disturbance treatments with a random effect for taxon (Eq. S1). This model structure avoids the use of interaction terms. Thus, the treatment effects from this model give the estimated differences between of the square root CVs of the ambient treatment, AA, and each of the 8 disturbed treatments.

$$
\begin{aligned}
\sqrt{CV} &= \beta_0 + \beta_1 X_{AD} + \beta_2 X_{AS} + \beta_3 X_{DA} + \beta_4 X_{SA} \\
&\quad + \beta_5 X_{DD} + \beta_6 X_{DS} + \beta_7 X_{SD} + \beta_8 X_{SS} \\
&\quad + \alpha_{taxon} + \varepsilon
\end{aligned}
\qquad \text{(Eq. S1)}
$$

$$\alpha_{taxon} \sim N(0, \sigma^2_{intercept})$$
$$\varepsilon \sim N(0, \sigma^2_{error})$$

*P values obtained from the theoretical distribution*

As in the main text, the p values obtained directly from the mixed model are the result of calculating the z score of the observed treatment effect in order to find the proportion of theoretical treatment effects that were more extreme than the observed treatment effect.

*P values obtained using bootstrapping*

In the bootstrapping approach, we obtained an empirical null distribution of treatment effects by randomizing the data before running the mixed model. We randomly assigned the 108 slides to the 9 treatments before calculating the square root CVs of the diatom taxa. In this approach, taxa remain associated by slide, but there is no true effect of treatment, because slides are randomly assigned to treatments. Then, we ran the mixed model on these square root CVs obtained from the randomized data and recorded the treatment estimates. Thus, the treatment estimates should include any effects of the interdependence of the taxa on the same slide. We repeated this workflow 1000 times to obtain a distribution of 1000 sets of treatment estimates. These 1000 sets of treatment estimates were then used as the null distribution of treatment effects, as they were all obtained under the condition where there was no real difference between treatments. Finally, we compared the treatment effects from the real, non-randomized dataset to the treatment effects obtained when slides were randomized into treatments. We obtained pseudo p values by calculating the proportion of treatment estimates in the null distribution that were more extreme (lower) than treatment estimates from the real dataset. This proportion gives the fraction of random slide assignments that led to a more negative estimated treatment effect than was observed in the true data.

We found that, similar to the mixed model, the overall trend was for populations to become more predictable (have a lower square root CV) after the experimental disturbances (Fig. S5). The associated pseudo p values are given in Table S1. The similarity of the results of this

analysis and the analysis presented in the main text suggests that the observed decrease in

population variability in disturbed treatments is a result that is robust to various statistical

methodologies.

**Table S1:** For each treatment, we calculated a pseudo p value, which was the fraction of the null

distribution of mixed model coefficients that was lower than the observed mixed model

coefficient. * indicates $p < 0.05$; ** indicates $p < 0.01$.

**Table S1: Comparison between theoretical and pseudo P values**

| Treatment | AD | AS | DA | DD | DS | SA | SD | SS |
|---|---|---|---|---|---|---|---|---|
| **Theoretical P-Value** | 0.0016 ** | 0.031 * | 0.033 * | 0.014 * | 0.0052 ** | 0.034 * | 0.0001 ** | 0.19 |
| **Pseudo P-Value** | 0.008 ** | 0.037 * | 0.044 * | 0.023 * | 0.006 ** | 0.038 * | 0.000 ** | 0.158 |

**Fig. S1:** At the start of the experiment, the 108 Plexiglas slides were distributed across 6 identical metal racks. All 6 racks were deployed side by side on buoy lines in Lake Myvatn and were suspended 0.3m from the sediment surface. After 20 days of periphyton colonization (T1), slides were again randomized before experiencing the first disturbance. Again, on day 25, slides were randomized onto new racks corresponding to the disturbance experienced at T2.

**Fig. S2**: We plotted the CVs of each diatom taxon within each treatment (72 total populations) against the mean abundance of that population. We found that the population CV was not biased by the mean population abundance.

**Fig. S3**: We plotted the estimated random effect from the mixed model against the log of the mean abundance of each diatom taxon. We found no relationship between the fitted random effects and the mean abundances of the diatom taxa. Abbreviations refer to *Cocconeis spp., Rhoicoshpenia spp., Cymbella spp., Synedra spp., Gomphonema spp., Nitzschia holsatica,* single *Fragilaria spp.*, and colonial *Fragilaria spp.*

**Fig. S4**: Mixed model residuals were not biased by treatment and were approximately normally distributed.

**Fig. S5:** We ran mixed models with our data randomly assigned to treatment to obtain a null distribution of treatment coefficients to compare to the true treatment coefficients. The mixed models estimated the effect of the 8 disturbance regimes on the square root CV of taxa within that treatment. Grey points show the 1000 coefficients obtained for each disturbance regime from mixed models run with samples that were randomly assigned to the 9 treatments. Black points show the observed coefficients (from the true data) from the same mixed model.

**Fig. S6:** A principal component analysis for the diatom communities shows that the AA treatment polygon overlaps strongly with every other treatment polygon. The AA treatment

polygon (red) spans a large portion of the first axis (PC 1), as well as the entire length of the second axis (PC 2).

**Fig. S7:** A principal component analysis for the bacterial communities shows that the AA treatment polygon (red) overlaps strongly with every other treatment polygon.

**Fig. S1**

**Fig. S2**
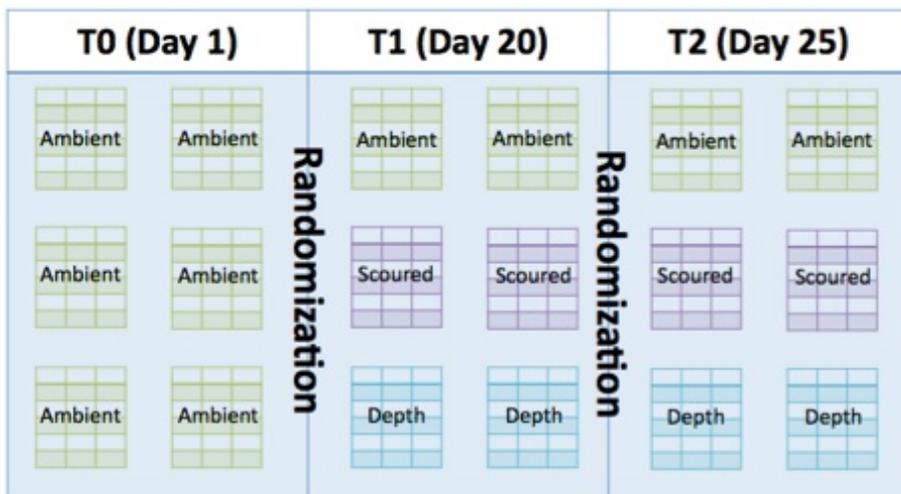
**Fig. S3**
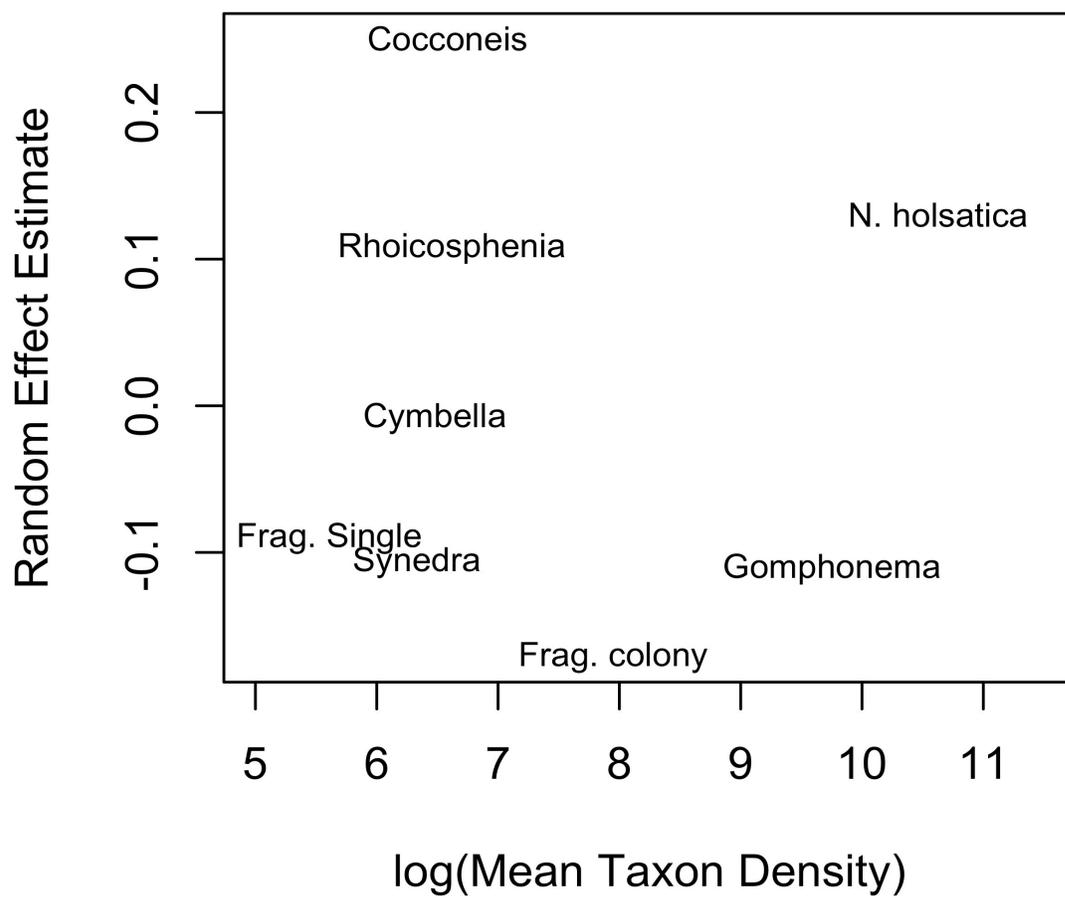
**Fig. S4**
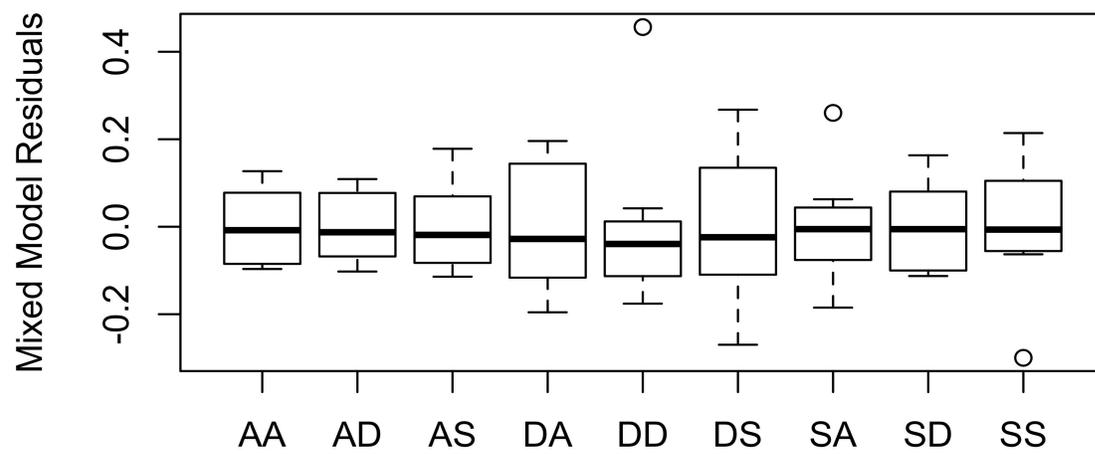
**Fig. S5**
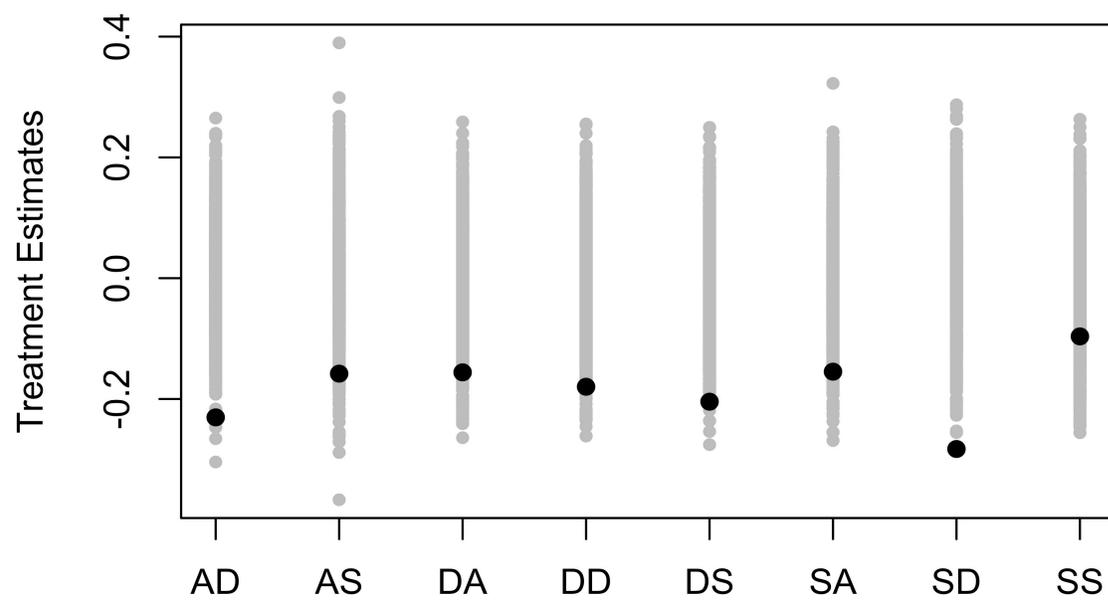
**Fig. S6**
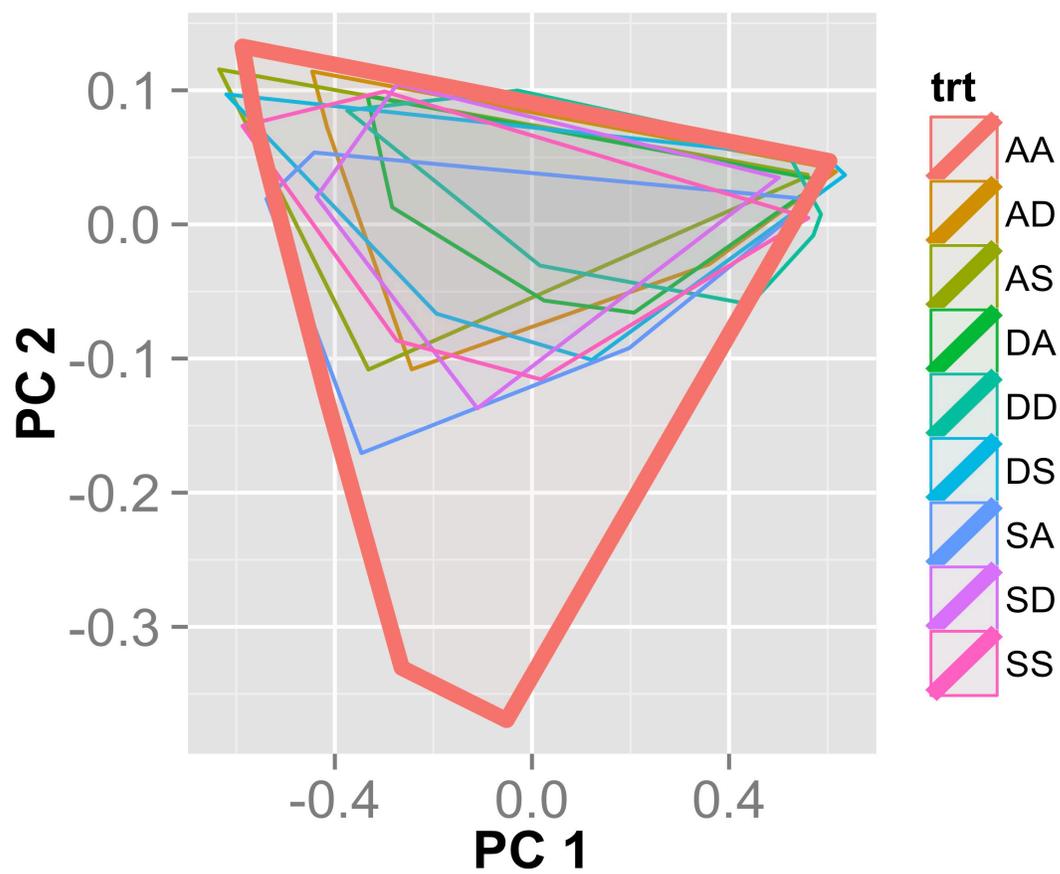
**Fig. S7**
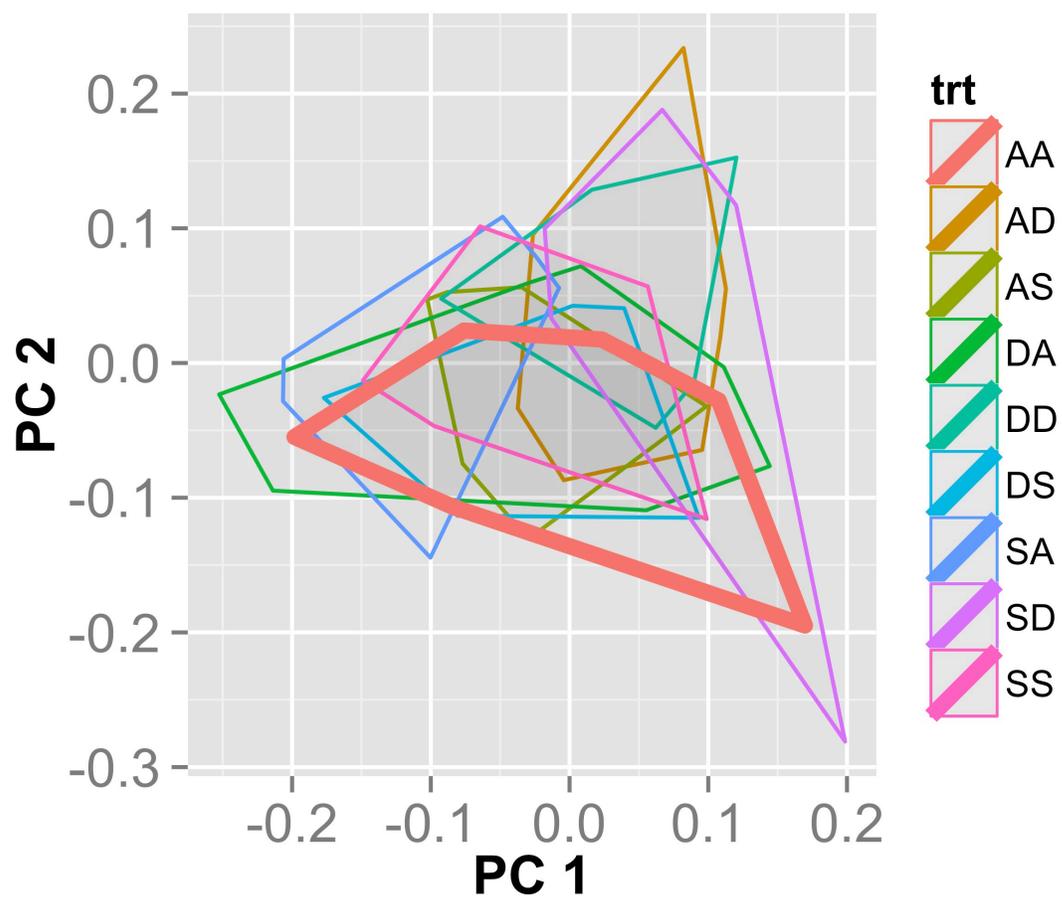
**Chapter 3:** Positive feedback between chironomids and algae creates net mutualism between benthic primary consumers and producers

Cristina M. Herren[1,2], Kyle C. Webert[2], Michael D. Drake[3], M. Jake Vander Zanden[4], Árni Einarsson[5], Anthony R. Ives[2], and Claudio Gratton[6]

[1]University of Wisconsin – Madison, Freshwater and Marine Science Program, Madison, WI, USA

[2]University of Wisconsin – Madison, Department of Zoology, Madison, WI, USA

[3]North Carolina State University, Department of Forestry and Environmental Resources, Raleigh, NC, USA

[4]University of Wisconsin – Madison, Center for Limnology, Madison, WI, USA

[5]Mývatn Research Station, Iceland and University of Iceland, Faculty of Life and Environmental Sciences, Reykjavík, Iceland

[6]University of Wisconsin – Madison, Department of Entomology, Madison, WI, USA

**Abstract:**

The chironomids of Lake Mývatn show extreme population fluctuations that affect most aspects of the lake ecosystem. During periods of high chironomid densities, chironomid larvae comprise over 90% of aquatic secondary production. Here, we show that chironomid larvae substantially stimulate benthic gross primary production (GPP) and net primary production (NPP), despite consuming benthic algae. Benthic GPP in experimental mesocosms with 140,000 larvae/m$^2$ was 71% higher than in mesocosms with no larvae. Similarly, chlorophyll *a* concentrations in mesocosms increased significantly over the range of larval densities. Furthermore, larvae showed increased growth rates at higher densities, possibly due to greater benthic algal availability in these treatments. We investigated the hypothesis that larvae promote benthic algal growth by alleviating nutrient limitation, and found that 1) larvae have the potential to cycle the entire yearly external loadings of nitrogen and phosphorus during the growing season and 2) chlorophyll *a* concentrations were significantly greater in close proximity to larvae (on larval tubes). The positive feedback between chironomid larvae and benthic algae generated a net mutualism between the primary consumer and primary producer trophic levels in the benthic ecosystem. Thus, our results give an example in which unexpected positive feedbacks can lead to both high primary and high secondary production.

**Introduction**

Predation is classically defined as an interaction where a consumer species exerts a net negative effect on a resource's per-capita growth rate, while receiving a net benefit from exploiting the resource population (Gotelli 2001). However, these long-term emergent effects on population growth rates are the result of many discrete, short-term interactions (Vázquez et al. 2015). During these short-term interactions, one species can exert either positive or negative effects on another species (Chamberlain and Holland 2009). For example, individuals from two species may experience a range of different interactions based on environmental factors (Juliano 2009), the abundance of other species (Paine 1969), or life stage (Pimm and Rice 1987). Thus, although consumers generally suppress resource abundance (Sih et al. 1985), they may also have short-term positive interactions with their resources.

Several studies have shown that consumers can have positive effects on their resources, often in the form of increased resource productivity (e.g. per-capita birth rate). Instances where this positive effect has been demonstrated include consumer-driven nutrient availability (McNaughton 1983, McIntyre et al. 2008, Knoll et al. 2009) compensatory plant growth following herbivore browsing (Petelle 1982), and in the context of optimum sustainable yield of fisheries (Beverton and Holt 1957). Examples where consumers might increase resource abundance are less common, but have been shown when predation shifts the resource age structure to have fewer adults and more juveniles (Zipkin et al. 2008). Still, empirical cases of positive feedbacks between consumers and food resources are scarce in comparison to the range of circumstances where they are predicted from theoretical models (Abrams 2009). Additionally, given the relatively little conceptual development of mutualistic interactions (Vázquez et al.

2015), it is unclear how this positive feedback within consumer-resource interactions might influence emergent properties such as population dynamics or ecosystem productivity.

Lake Mývatn is an anomaly for a subarctic lake, with an abundant and diverse food web (Einarsson et al. 2004). Despite a short growing season and cool temperatures, the lake has surprisingly high primary and secondary productivity for its high latitude (Lindegaard and Jónasson 1979). The most striking example of this high productivity are the chironomids (Diptera: Chironomidae) that can occur at densities $> 500,000/m^2$ in the benthos (Thorbergsdóttir et al. 2004). Chironomid populations at Mývatn fluctuate over 3-4 orders of magnitude but routinely reach very high densities, comprising $> 90\%$ of secondary production (Einarsson et al. 2002, Ives et al. 2008). It is evident that, in supporting this amount of chironomid biomass, algal growth rates are able to keep pace with the grazing pressure of these primary consumers.

We hypothesize that chironomid larvae, in addition to being consumers of the benthic algae in Lake Mývatn, also have a strong positive effect on algal productivity. We hypothesize that primary production and algal biomass increase in response to high larval densities, creating a positive feedback within this consumer-resource system. Furthermore, we hypothesize that the positive feedback might be sufficiently strong as to increase secondary production due to a higher short-term growth rate of chironomid larvae (i.e. body size increases) as their resource availability increases. We investigate two related mechanisms that could lead to this positive effect of primary consumers on primary producers: larval tubes provide a superior substrate for benthic algal growth (Pringle 1985), and larval excretion increases the availability of limiting nutrients (Atkinson et al. 2013). We hypothesize that the positive effect of chironomid larvae on their food resources is one factor that contributes to the high primary and secondary productivity of the Lake Mývatn system.

**Methods**

*Lake Mesocosms Across a Range of Larval Densities*

If chironomid larvae facilitate algal production, then we would expect that as larval densities increase, so would benthic gross primary production (hereafter, benthic GPP). This experiment was designed to determine how varying densities of chironomid larvae affect benthic algal productivity, chlorophyll *a* concentration, and growth rates of chironomid larvae. We collected chironomid larvae and sediment from Lake Mývatn using an Ekman grab on 11-14 July 2014. We identified and sorted 23,450 live chironomid larvae for this experiment. Larvae were identified to tribe, and chironomini and tanytarsini were collected for the experiment. At the collection location, almost all chironomini were *Chironomus islandicus* (Kieffer) larvae, and almost all of the tanytarsini were *Tanytarsus gracilentus* (Holmgren) larvae; therefore, we will refer to the chironomini as *C. islandicus* and the tanytarsini as *T. gracilentus*.

On 15 July 2014, we set up 55 mesocosms stocked with sieved (125 µm) lake sediments and with 8 levels of chironomid densities. Mesocosms consisted of 1-L clear polypropylene deli cups (10.4 cm diameter and 16 cm height) filled to a depth of 10 cm with sieved sediment and left uncovered at the top. We filled mesocosms with assemblages of 75% *C. islandicus* larvae and 25% *T. gracilentus* larvae, as that was the ratio of larvae recovered from Ekman grabs. These two taxa comprised the overwhelming majority (>95%) of organisms recovered from Ekman grabs. Both *C. islandicus* and *T. gracilentus* are vertical tube-building chironomids that feed non-selectively on detritus and diatoms present at the sediment surface and on their larval tubes (Einarsson et al. 2004). The two species are primarily differentiated in their ecology by size (*C. islandicus* is the larger species, averaging 9 times the mass of *T. gracilentus* larvae during this study) and depth of burrows; *T. gracilentus* often builds tubes that are 2-3 cm in

length near the sediment surface (Ólafsson and Paterson 2004), but we often observed *C. islandicus* building vertical burrows that extended to between 10-15 cm depth. The bioturbation activities of both species include ventilation of their blind-end burrows and particle reworking that results in biodiffusion (sensu Kristensen et al. 2012). We used the following larval densities for experimental mesocosms: 0, 50, 100, 200, 400, 600, 800, and 1200 per mesocosm. These numbers correspond to the following densities of larvae/m$^2$: 0, 5 886, 11 772, 23 544, 47 087, 70 631, 94 174, and 141 262. The amount of biomass introduced in the mesocosms with 1200 larvae was intended to exceed the maximum larval biomass that has been observed in lake sediments (Thorbergsdottir et al. 2004). There were six replicates of the treatment with 0 larvae and seven replicates of all other treatments. We randomly distributed the mesocosms across 11 metal racks and set them on the lake bottom in an area of the lake that was 3.5 m deep, almost the maximum natural lake depth.

On 27 July 2014 (12 days later), we retrieved the mesocosms and performed incubations to analyze benthic GPP and benthic NPP of algae in the mesocosm sediments. Two mesocosms of each treatment were incubated in dark conditions (by wrapping mesocosms with shading tarp), and the remaining mesocosms were incubated under light conditions, which consisted of placing racks at a depth of 1-2 m in the lake to prevent light limitation of photosynthesis. During the incubations, we sealed mesocosms with Parafilm for a duration of 4 hours. We measured the dissolved oxygen concentration in the water column of each mesocosm (mean height and volume of water column = 6.3cm, 535mL) before and after the sealed incubation using a YSI ProODO probe (Yellow Springs, Ohio, USA). We gently mixed the water column with the probe to homogenize any potential vertical gradients in dissolved oxygen. Benthic NPP was measured as the net oxygen change in light-incubated mesocosms. Benthic GPP was estimated using the net

change in oxygen concentrations in light-incubated mesocosms plus the average oxygen consumption in the dark-incubated mesocosms of the same larval density (Vander Zanden et al. 2006). More of the mesocosms were incubated under light conditions than dark conditions because pilot studies indicated that light-incubated mesocosms had more variable oxygen measurements than dark mesocosms. We used linear mixed models (using fixed and random effects) to analyze benthic GPP and NPP of the mesocosms. The fixed effect was the number of larvae in the mesocosms, and the random effect was the rack where the mesocosm was situated. The fixed effect was the number of larvae in the mesocosms, and the random effect was the rack where the mesocosm was situated.

Immediately after benthic GPP incubations were completed, we took sediment samples of the top 0-1 cm of the mesocosms and froze them at -20 ºC. Within 3 weeks, sediments were analyzed for chlorophyll *a* content. We analyzed chlorophyll *a* concentration by extracting 1 mL of sediment in methanol for 24 hours before quantifying fluorescence using standard protocols (Welschmeyer 1994) with a tabletop fluorometer (Turner Designs, Sunnyvale, CA, USA). We used a linear model to analyze the chlorophyll *a* content of the mesocosms as a function of the number of larvae in the mesocosms. Additionally, we used a linear model to analyze how chlorophyll-specific GPP (defined as the GPP of a mesocosm divided by the chlorophyll *a* concentration in the mesocosm) changed in response to larval density.

After chlorophyll *a* samples were taken, we then collected the chironomid larvae from the mesocosms to obtain the average dry weight of larvae in each mesocosm. The majority of remaining larvae were *C. islandicus*, because most *T. gracilentus* had pupated and emerged during the experiment. We placed these larvae in tap water at 4 ºC for 36 hours to allow them to void their gut contents. Then, we haphazardly sampled 30 *C. islandicus* larvae from each

mesocosm (or, all remaining *C. islandicus*, in the case of some of the mesocosms beginning with 50 larvae) and dried the larvae at 60 ºC for a minimum of 24 hours to obtain an average larval dry weight for each mesocosm (Dermott and Paterson 1974). Prior to the beginning of this experiment, we had weighed a random subsample of 100 *C. islandicus* individuals to estimate initial larval weight. We used a linear model to analyze the dry weight of *C. islandicus* larvae as a function of the number of larvae stocked in the mesocosms.

*Lake Mesocosms for Larval Tubes*

We hypothesized that the silken protective tubes spun by the larvae would contain elevated levels of chlorophyll *a*, as compared to loose sediments, because of the high substrate quality of larval tubes (Pringle 1985). On 12 July 2014, we used similar protocols as described above (larval density experiment) to establish mesocosms with chironomid larvae separated to species. These mesocosms were constructed using 400 mL containers (7.0 cm diameter) filled with approximately 275 mL of sieved sediment to a depth of 8 cm. We filled 8 mesocosms with 175 larvae of *Chironomus islandicus*, 8 mesocosms with 175 larvae of *Tanytarsus gracilentus,* and 10 mesocosms with no larvae. We placed mesocosms on racks and submerged them at 3.5 m depth in Lake Mývatn. On 21 July 2014 (9 days), we retrieved the mesocosms and sampled the top 1cm of all mesocosms for chlorophyll *a* content. We sampled chironomid tubes from the two species by collecting 0.5 mL of larval tubes. To collect the tubes, we used forceps to remove adjacent sediments and extracted the top 1 cm of the tube. We compared chlorophyll *a* content in larval tubes of each species to the chlorophyll *a* content in adjacent loose sediments.

*Quantification of Larval Excretion*

A hypothesis regarding the mechanism of the positive feedback between chironomid larvae and benthic algae was that the larvae increase nutrient availability to the algae. To

investigate the magnitude of the soluble nutrient flux, we conducted incubations of chironomid larvae to quantify the amount of nitrogen and phosphorus in larval excretions.

We collected larvae with Ekman grabs on 22 July 2014. We transported the larvae back to the lab, where groups of 100 larvae were collected and immediately placed into 100 mL of distilled water in new, amber Nalgene bottles. We incubated these larvae in the shade outside for 4 hours. Temperatures during the incubations were between 15.8-18.0 ºC, which is within the range of observed summer water temperatures in Lake Mývatn. Immediately afterwards, we sieved the contents of the bottles through 63 μm mesh to remove chironomid larvae and any fecal material they had passed before obtaining larval dry weight (Dermott and Paterson 1974). The water samples from these incubation experiments were then frozen at -20 ºC and were transported on dry ice to Madison, Wisconsin, USA, where samples were analyzed for soluble nitrogen (combined $NH_4$, $NO_3^-$, $NO_2^-$) and soluble reactive phosphorus (SRP) following protocols used by the North Temperate Lakes Long Term Ecological Research program (www.lter.limnology.wisc.edu). All statistical analyses in this study were performed with the R programming environment (v. 3.1.3), using the *base* and *stats* packages for data handling and linear regression, the *lme4* package for linear mixed models, and the *outliers* packages for detecting anomalous data points.

**Results**

*Lake Mesocosms Across a Gradient of Larval Densities*

Benthic GPP in mesocosms responded strongly and positively to the presence of chironomid larvae (n = 39, t = 17.8; p < 0.001, conditional $R^2$ = 0.93, Fig. 1a). Furthermore, there was also an increase in NPP with increasing numbers of chironomid larvae (n = 39, t = 7.64; p < 0.001, Fig. 1b). Thus, there were greater absolute oxygen concentrations in mesocosms with high

densities of chironomid larvae, despite the increased oxygen consumption by microbes that are stimulated by larval activity and, to a lesser extent, respiration of the larvae themselves (Baranov et al. 2016). We assumed that mortality of larvae among treatments was small and not substantially different by treatment, as differences in larval densities across treatments were still visually obvious when the mesocosms were retrieved.

Chlorophyll *a* concentrations increased linearly in response to higher densities of chironomid larvae ($F_{1, 53}$ = 15.4, p < 0.001, Fig. 1c). For every 100 larvae added to a mesocosm, chlorophyll *a* concentrations increased by 1.5% relative to mesocosms with no larvae, with no apparent saturation over the densities used in this study.

The average individual mass of *C. islandicus* larvae increased significantly in response to higher densities of larvae stocked in the mesocosms ($F_{1, 47}$ = 8.23, p = 0.0062). That is, larvae stocked at high densities had grown more than larvae stocked at lower densities (Fig. 1d). Furthermore, average mass of larvae stocked at low densities decreased from the average starting mass ($\Delta$ dry weight = - 0.052 mg for the lowest density of 50 larvae), whereas larvae stocked at high densities increased in average mass ($\Delta$ dry weight = + 0.053 mg DW for the highest density of 1200 larvae, approximately + 6% increase).

Chlorophyll-specific GPP strongly increased ($F_{1, 37}$ = 53.0, p < 0.001) in response to larval density (Fig. 1e). Thus, benthic GPP increased at a faster rate than chlorophyll *a* concentration with increasing larval densities.

*Lake Mesocosms for Larval Tubes*

Chironomid tubes of both species had more than double the chlorophyll *a* concentrations of adjacent loose sediments. For mesocosms containing *T. gracilentus,* larval tubes had significantly greater chlorophyll *a* concentrations (mean ± standard error = 51.2 mg/L ± 2.63

mg/L), than loose sediments (mean = 22.8 mg/L ± 1.75 mg/L, $F_{1, 16}$ = 117, p < 0.001). Similarly, *C. islandicus* tubes had a mean chlorophyll *a* concentration of 47.0 mg/L (± 2.38 mg/L) and sediments had 21.8 mg/L (±1.58 mg/L, $F_{1, 16}$ = 112, p < 0.001). However, tubes of the two species did not have significantly different chlorophyll *a* concentrations ($F_{1, 14}$ = 1.38, p = 0.26).

*Quantification of Larval Excretion*

We obtained measurements of soluble nitrogen (combined $NH_4^+$, $NO_3^-$, and $NO_2^-$) and soluble reactive phosphorus (SRP) for five incubations of *T. gracilentus* and four incubations of *C. islandicus,* after removing 1 outlier sample. Of the soluble N, 94% was in the form of $NH_4^+$. For soluble N, *C. islandicus* individuals excreted 1.05 μg/d (± 0.19 SEM), whereas *T. gracilentus* individuals excreted 0.38 μg/d (± 0.17). However, when standardized by dry weight, *C. islandicus* excreted 1.54 $μg \cdot d^{-1} \cdot mg^{-1}$ (± 0.82) and *T. gracilentus* excreted 4.94 $μg \cdot d^{-1} \cdot mg^{-1}$ (± 0.73). For SRP, these values were 0.26 μg/d (± 0.020) for *C. islandicus* individuals and 0.10 μg/d (± 0.018) for *T. gracilentus* individuals. By dry weight, the equivalent excretion rates were 0.37 $μg \cdot d^{-1} \cdot mg^{-1}$ (± 0.055) for *C. islandicus* and 1.32 $μg \cdot d^{-1} \cdot mg^{-1}$ (± 0.050) for *T. gracilentus*.

**Discussion**

Our results indicate that chironomid larvae can generate a net positive effect on their algal resources. Despite the fact that benthic algae are a primary food source for chironomid larvae (Ingvason et al. 2004), benthic algae were more abundant and more productive with higher numbers of chironomid larvae stocked in our mesocosms. In mesocosms with 1 200 larvae, benthic GPP was, on average, 71% greater than in mesocosms with no larvae. Additionally, benthic algal biomass (assessed by chlorophyll *a*) was greater in mesocosms that had been stocked with more chironomid larvae, despite the ongoing grazing of algae by the larvae. More critically, however, benthic GPP increased faster than chlorophyll *a*, indicating that

productivity per unit of algal biomass increased as more chironomid larvae were present. Furthermore, *C. islandicus* larvae also grew faster when stocked at higher densities; the increased mass of *C. islandicus* larvae at high densities suggests that the consumer-resource positive feedback is sufficiently strong to generate positive density-dependence of the growth of chironomid larvae. Thus, these experiments show that, at high larval densities, chironomid resource limitation was alleviated by their stimulation of benthic algae. Furthermore, Fig. 1e shows that chlorophyll-specific GPP increased at high consumer biomass; similarly, Fig. 1d shows that *C. islandicus* also reached greater individual mass at high densities. This chironomid-algal positive feedback is one potential mechanism that may explain how the chironomids can reach such extreme densities in the benthos of Lake Mývatn.

*Magnitude of consumer-driven nutrient cycling*

Alleviation of nutrient limitation by larval activity could play a role in generating the observed positive feedback, especially because internal nutrient cycling is the largest contributor to the nitrogen and phosphorus budgets in Lake Mývatn (Ólafsson 1979) and many other freshwater systems (Vanni 2002). We found that larval excretions provide concentrated nutrients in the close proximity of benthic algae, which could account for the high quality of larval tubes as a substrate for algae (Hershey et al. 1988). Another possible positive effect of tube building on algal growth rates is the change in physical structure the tubes create. The three-dimensional silk tubes of *C. islandicus* and *T. gracilentus* larvae increase surface area for algal growth, while altering physical characteristics of surface sediments due to the binding properties of larval silk (Ólafsson and Paterson, 2004). Furthermore, because the chironomid larvae also consume detritus as 30-50% of their diet (Einarsson et al. 2004, Ingvason et al. 2004), the larvae also mobilize nutrients stored in decaying organic matter and diatom fragments. Thus, the larvae

likely increase nutrient availability by both increasing the rate of nutrient cycling and by adding to the pool of biologically available nitrogen and phosphorus through their digestion of detritus.

For illustrative purposes, we can examine the potential significance of nutrient mobilization by larvae relative to other sources in the lake. By multiplying chironomid excretion rates (for nitrogen, 1.05 and 0.38 µg/d for *C. islandicus* and *T. gracilentus*, respectively) by an estimate of larval density from 19 Aug 2014 from one location in the lake (7,500 *C. islandicus* per $m^2$ and 431,000 *T. gracilentus*), we find that chironomids could move in 8 days the equivalent of the yearly external input of nitrogen (1.4 g N/$m^2$ $y^{-1}$, Ólafsson 1979), which is often the limiting nutrient in the lake (Ólafsson 1979, Einarsson et al. 2004). We can similarly estimate that chironomid larvae could cycle the yearly input of phosphorus (1.5 g P/$m^2$ $y^{-1}$) in 30 days. This is based on several simplifying assumptions, including constant excretion and flux rates, as estimated from the excretion assay, that chironomid densities (and size structure) remain relatively constant, and that excretion was the sole source of nutrients in the incubation water samples. Still, these estimates further demonstrate the potential for chironomids to influence whole-ecosystem processes (Hölker et al. 2015). Similarly, prior research has found that chironomids may have a substantial impact on benthic nutrient cycling above densities of 1 000 individuals/$m^2$ (Tátrai 1988). Additionally, chironomid behavior may also affect nutrient availability in the benthos, because chironomid bioturbation stimulates $NH_4^+$ release from sediments (Tátrai 1988, Hölker et al. 2015). In a prior study, another *Chironomus* species liberated $NH_4^+$ from sediments down to at least 15 cm depth (Lewandowski et al. 2007). In Lake Mývatn, $NH_4^+$ concentrations are very high in deep sediments, reaching > 1 000 µg/L by 20 cm depth (Gíslason et al. 2004). These results also suggest that the rates of nutrient cycling in Lake Mývatn could vary dramatically from year to year, based on larval densities.

*Implications for rates of ecosystem productivity*

The magnitude of the consumer-resource positive feedback suggests that variability in chironomid abundance could be a strong determinant of primary production in Lake Mývatn. Additionally, the net mutualism in this consumer-resource system is sufficiently strong as to increase secondary production. It is evident that the positive feedback can generate positive density-dependence in the chironomids, as the *C. islandicus* larvae grew more quickly in mesocosms where they were at high densities. One hypothesis to explain the positive density-dependence in *C. islandicus* is that larvae feed both on and around their tubes, meaning that these chironomids could increase algal availability to neighboring larvae if individuals are sufficiently close. Previous studies of chironomid larval behavior have found evidence of spatial aggregation (Titmus and Badcock 1981, Drake 1983), and we propose that the positive feedback observed in this study may promote behavioral aggregation due to greater local resource availability. Interestingly, the chironomids from treatments with comparatively low larval densities lost mass during the experiment. We hypothesize that all larvae in our mesocosms likely experienced a decrease in their resource availability as a result of establishment of mesocosms. Homogenizing the sediments would have reduced the algal availability in mesocosms, as compared to the original lake sediments. This is because it disrupted the high-quality periphyton resources that were present on sediment surfaces and on larval tubes. Furthermore, larvae could have lost body mass due to metabolic costs incurred as a result of building new tubes and burrows in the mesocosms. For these reasons, we expect that the mass of all larvae would have declined initially upon being relocated to mesocosms. However, larvae in higher density treatments may have regained body mass more quickly due to the higher primary production of these mesocosms.

We also hypothesize that mutualisms between consumers and resources, such as the positive feedback described here, could ultimately translate to population-level effects in these systems. For example, for chironomids at Lake Myvatn, this positive feedback might explain the prolonged period of exponential growth experienced by the chironomids during the upswings in their population fluctuations; the mutualism between the chironomid consumers and the algal resources may be sufficiently strong as to allow for several generations of exponential growth due to a lack of resource limitation. However, it is obvious that at some point, the positive feedback between the chironomid community and the benthic algae breaks down, because the chironomid populations eventually crash. As such, we hypothesize that the relationship between density of chironomid larvae and the rate of benthic algal production may be dependent on other factors not considered in this study, which could lead to a non-linear relationship under different conditions. For example, the proportion of detritus in sediments could be an important modulating factor in the positive feedback (de Mazancourt et al. 1998), as it represents a source of nutrients that are added to the nutrient pool, rather than only recycled within the nutrient pool. Additionally, the stage structure of the chironomid larvae could be another factor determining the strength of this mutualism. The biomass of larvae in the lake changes substantially throughout the growing season as a result of larvae progressing through stages of development, which could alter the relative rates of consumption and stimulation of benthic algae. Because the chironomids in Lake Mývatn have a finite population size, the linear relationship between chironomid density and chironomid growth must plateau under some circumstances.

Although we have highlighted several reasons to believe that the positive relationship between chironomids and their food resources must break down under different conditions in our study system, it is worthwhile to consider the theoretical case where this effect is uniform across

space and time. In the case that all else is held constant, a positive feedback that leads to differential consumer growth rates would be expected to influence population dynamics. Integrating positive feedbacks in consumer-resource systems with the existing theoretical frameworks of consumer-resource population dynamics suggests that systems with consumer-resource mutualisms may have specific population-level characteristics. For example, increased population-level growth rates in the consumer as a result of a mutualism could be considered a type of Allee effect (Allee 1931). Following this framework, a region of positive density-dependence in the consumer should generate a stable equilibrium at a higher density than would occur without the positive feedback (Stephens et al. 1999). In the context of the Lake Mývatn system, this means that the lake benthos might exhibit a tendency to support higher densities of larvae than would occur in the absence of the positive feedback. Furthermore, positive feedbacks within consumer-resource systems often tend to increase the birth rate of the resource (Bianchi et al. 1989). This scenario of variable resource birth rates has been long-studied by theoretical ecologists because it generates the "Paradox of Enrichment," which occurs when an increase in the per-capita birth rate of a resource destabilizes the population dynamics of a consumer-resource system (Rosenzweig 1971). This well-known result suggests that a positive feedback in a consumer-resource system might predispose the system to instability or cyclic dynamics. Thus, including positive feedbacks into models of consumer-resource systems may have predictable consequences for population dynamics.

*Positive feedbacks in consumer-resource interactions*

Although chironomid larvae could generate a net positive effect on primary production, these larvae are still consumers of the benthic algae. Thus, the interaction between the larvae and the benthic algae was a combination of facilitation and consumption. Because the larvae

stimulated algal growth to a greater extent than they suppressed it through consumption, the larvae and the algal showed a net mutualistic interaction, whereby both taxa had a greater growth rate when the other was highly abundant. Whereas combinations of species interactions such as competition and predation have been considered simultaneously (e.g. intraguild predation), a similar framework for understanding mutualism in the context of predation is lacking (Vázquez et al. 2015). This study suggests that strong consumer-resource positive feedbacks can affect ecosystem-level properties, such as primary production and producer-to-consumer ratios.

This consumer-resource positive feedback may be important in other systems where grazers are the primary consumers. We investigated here two mechanisms that may contribute to the positive effects of grazers on their resources: nutrient cycling through excretion and ecosystem engineering providing enhanced substrate for growth. However, another contributing factor here might be that benthic bioturbators increase the flux of soluble nutrients out of sediment pore water, making these nutrients available to primary producers (Lewandowski & Hupfer 2005). More generally, there are several other mechanisms whereby grazers could promote the growth of their resources. These mechanisms include increased growth due to a release from shading following grazing, hormone secretions in grazer saliva that induce growth, removal of senescing tissue, release from intraspecific competition, and greater efficiency in resource distribution within the remaining population (McNaughton 1983). Given the myriad mechanisms by which grazers could increase the growth rate of their resources, it would be interesting to look for this positive feedback in other systems, especially investigating whether background nutrient availability influences the strength of the positive feedback.

A recent review of the effect of tube-dwelling invertebrates as ecosystem engineers concluded that it is difficult to determine the importance of tube-dwelling invertebrates at the

lake-wide scale (Hölker et al. 2015). However, this study supports the hypothesis that chironomids can be drivers of lake primary production by altering critical feedback loops. This study demonstrates that the positive feedback between tube-dwelling invertebrates and benthic primary producers can significantly affect lake-wide productivity, as benthic GPP can account for upwards of 80% of total aquatic GPP in this system (Jónasson 1979, Thorbergsdóttir et al. 2004). Although the effects of tube-dwelling invertebrates on pelagic primary production have been well studied (Hölker et al. 2015), the critical role these ecosystem engineers play in the benthos has not been documented as thoroughly, even though benthic primary production may be greater than pelagic primary production in shallow lakes (Vadeboncoeur et al. 2008). Finally, this positive feedback may be particularly important in regions where chironomids are used as indicators of water quality, if the chironomid larvae modify their environment by stimulating algal production.

Several previous studies have shown that consumers can have marginal positive effects on resources (McNaughton 1983, Holland and DeAngelis 2010), even if the net effect of consumers is generally negative (Sih et al. 1985). Despite this observation, positive feedbacks have often been ignored in conceptual models of consumer-resource interactions (Vázquez et al. 2015). However, in this instance, the consumer-resource facilitation was stronger than the effect of consumption; the benthic primary consumer trophic level exerted a net positive effect on their resources. Thus, contrary to the paradigm that aquatic primary consumers are particularly effective at suppressing primary producer biomass (Hairston and Hairston 1993, Shurin et al. 2006), we found that increasing the density of primary consumers could substantially increase benthic NPP and primary producer biomass. These results also highlight the importance of considering the rates of turnover of resources, in addition to standing biomass. Facilitation in

consumer-resource interactions, which is often overlooked when net interaction strengths are negative, has the potential to substantially alter traditional consumer-resource dynamics. In the Mývatn system, positive consumer-resource feedbacks can be sufficiently strong as to create a net mutualism between the primary producer and primary consumer trophic levels in the benthos.

**Acknowledgements**

**References**

Abrams, P. A. 2009. When does greater mortality increase population size? The long history and diverse mechanisms underlying the hydra effect. Ecology Letters 12.5:462-474.

Allee, W. C. 1931. Animal aggregations, a study on general sociology. University of Chicago Press, Chicago, Illinois, USA.

Atkinson, C. L., C. C. Vaughn, K. J. Forshay, and J. T. Cooper. 2013. Aggregated filter-feeding consumers alter nutrient limitation: consequences for ecosystem and community dynamics. Ecology 94:1359–1369.

Baranov, V., J. Lewandowski, P. Romeijn, G. Singer, and S. Krause. 2016. Effects of bioirrigation of non-biting midges (Diptera: Chironomidae) on lake sediment respiration. Scientific Reports 6:27329.

Beverton, R. J. H., and S. J. Holt. 1957. On the dynamics of exploited fish populations. Fisheries Investigation Series 2(19). Ministry of Agriculture, Fisheries, and Food, London, UK.

Bianchi, T. S., C. G. Jones, and M. Shachak. 1989. Positive feedback of consumer population density on resource supply. Trends in Ecology & Evolution 4:234–238.

Chamberlain, S. A., and J. N. Holland. 2009. Quantitative synthesis of context dependency in ant-plant protection mutualisms. Ecology 90:2384–2392.

de Mazancourt, C., M. Loreau, and L. Abbadie. 1998. Grazing optimization and nutrient cycling: when do herbivores enhance plant production? Ecology 79:2242–2252.

Dermott, R. M., and C. G. Paterson. 1974. Determining dry weight and percentage dry matter of chironomid larvae. Canadian Journal of Zoology 52:1243–1250.

Drake, C. M. 1983. Spatial distribution of chironomid larvae (Diptera) on leaves of the bulrush in a chalk stream. Journal of Animal Ecology 52:421–437.

Einarsson Á., A. Gardarsson, G. M. Gíslason, and A. R. Ives. 2002. Consumer-resource

interactions and cyclic population dynamics of Tanytarsus gracilentus (Diptera: Chironomidae). Journal of Animal Ecology 71:832-845.

Einarsson, Á., G. Stefánsdóttir, H. Jóhannesson, J. S. Ólafsson, G. M. Gíslason, I. Wakana, G. Gudbergsson, and A. Gardarsson. 2004. The ecology of Lake Myvatn and the River Laxá: Variation in space and time. Aquatic Ecology 38:317–348.

Gíslason, S. R., E. S. Eiríksdóttir, and J. S. Ólafsson. 2004. Chemical composition of interstitial water and diffusive fluxes within the diatomaceous sediment in Lake Myvatn, Iceland. Aquatic Ecology 38:163–175.

Gotelli N. J. 2001. A Primer of Ecology. Third edition. Sinauer Associates. Sunderland, Massachusetts, USA.

Hairston, N. G. Jr. and N. G. Hairston Sr. 1993. Cause-effect relationships in energy flow, trophic structure, and interspecific interactions. The American Naturalist 142:379–411.

Hershey, A. E., A. L. Hiltner, M. A. J. Hullar, M. C. Miller, J. R. Vestal, M. A. Lock, S. Rundle, and B. J. Peterson. 1988. Nutrient incluence on a stream grazer: Orthocladius microcommunities respond to nutrient input. Ecology 69:1383–1392.

Hölker, F., M. J. Vanni, J. J. Kuiper, C. Meile, H.-P. Grossart, P. Stief, R. Adrian, A. Lorke, O. Dellwig, A. Brand, M. Hupfer, W. M. Mooij, G. Nützmann, and J. Lewandowski. 2015. Tube-dwelling invertebrates: tiny ecosystem engineers have large effects in lake ecosystems. Ecological Monographs 85:333–351.

Holland, J. N., and D. L. DeAngelis. 2010. A consumer–resource approach to the density-dependent population dynamics of mutualism. Ecology 91:1286–1295.

Ingvason H., J. S. Ólafsson, and A. Gardarsson. 2004. Food selection of *Tanytarsus gracilentus* larvae (Diptera: Chironomidae): An analysis of instars and cohorts. Aquatic Ecology 38: 231-237.

Ives, A. R., Á. Einarsson, V. A. A. Jansen, and A. Gardarsson. 2008. High-amplitude fluctuations and alternative dynamical states of midges in Lake Myvatn. Nature 452:84–87.

Jónasson, P. M. 1979. The Lake Mývatn ecosystem, Iceland. Oikos 32:289–305.

Juliano, S. A. 2009. Species interactions among larval mosquitoes: context dependence across habitat gradients. Annual review of entomology 54:37–56.

Knoll, L. B., P. B. McIntyre, M. J. Vanni, and A. S. Flecker. 2009. Feedbacks of consumer nutrient recycling on producer biomass and stoichiometry: separating direct and indirect effects. Oikos 118:1732–1742.

Kristensen, E., G. PenhaLopes, M. Delefosse, T. Valdemarsen, C. O. Quintana, and G. T. Banta. 2012. What is bioturbation? The need for a precise definition for fauna in aquatic sciences. Marine Ecology Progress Series 446:285–302.

Lewandowski, J., C. Laskov, and M. Hupfer. 2007. The relationship between Chironomus plumosus burrows and the spatial distribution of pore-water phosphate, iron and ammonium in lake sediments. Freshwater Biology 52:331–343.

Lewandowski, J., and M. Hupfer. 2005. Effect of macrozoobenthos on two-dimensional small-scale heterogeneity of pore water phosphorus concentrations in lake sediments: A laboratory study. Limnology and Oceanography, 50.4:1106–1118.

Lindegaard, C. and P. M. Jónasson. 1979. Abundance, population dynamics and production of zoobenthos in Lake Mývatn, Iceland. Oikos 32:202–227.

McIntyre, P. B., A. S. Flecker, M. J. Vanni, J. M. Hood, B. W. Taylor, and S. A. Thomas. 2008. Fish distributions and nutrient cycling in streams: can fish create biogeochemical hotspots. Ecology 89:2335–2346.

McNaughton, S. J. 1983. Compensatory plant growth as a response to herbivory. Oikos 40:329–

336.

Ólafsson, J. 1979. The chemistry of Lake Mývatn and River Laxá. Oikos 32:82–112.

Ólafsson, J. S., and D. M. Paterson. 2004. Alteration of biogenic structure and physical properties by tube-building chironomid larvae in cohesive sediments. Aquatic Ecology 38:219–229.

Paine, R. T. 1969. A note on trophic complexity and community stability. The American Naturalist 103:91–93.

Petelle, M. 1982. More mutualisms between consumers and plants. Oikos 38:125–127.

Pimm, S. L., and J. C. Rice. 1987. The dynamics of multispecies, multi-life-stage models of aquatic food webs. Theoretical Population Biology 32:303–325.

Pringle, C. M. 1985. Effects of Chironomid (insecta: Diptera) tube-building activities on stream diatom communities1,2. Journal of Phycology 21:185–194.

Rosenzweig M. L. 1971. Paradox of enrichment: destabilization of exploitation ecosystems in ecological time. Science. 171:385–387.

Shurin, J. B., D. S. Gruner, and H. Hillebrand. 2006. All wet or dried up? Real differences between aquatic and terrestrial food webs. Proceedings of the Royal Society of London B: Biological Sciences 273:1–9.

Sih, A., P. Crowley, M. McPeek, J. Petranka, and K. Strohmeier. 1985. Predation, competition, and prey communities: a review of field experiments. Annual Review of Ecology and Systematics 16:269–311.

Stephens, P. A., W. J. Sutherland, and R. P. Freckleton. 1999. What is the Allee Effect? Oikos 87:185–190.

Tátrai, I. 1988. Experiments on nitrogen and phosphorus release by Chironomus plumosus from the sediments of Lake Balaton, Hungary. Internationale Revue der gesamten Hydrobiologie

und Hydrographie 73:627–640.

Thorbergsdóttir, I. M., S. R. Gíslason, H. R. Ingvason, and Á. Einarsson. 2004. Benthic oxygen

flux in the highly productive subarctic Lake Myvatn, Iceland: In situ benthic flux chamber

study. Aquatic Ecology 38:177–189.

Titmus, G., and R. M. Badcock. 1981. Distribution and feeding of larval Chironomidae in a

gravel-pit lake. Freshwater Biology 11:263–271.

Vadeboncoeur, Y., G. Peterson, M. J. Vander Zanden, and J. Kalff. 2008. Benthic algal

production across lake size gradients: interactions among morphometry, nutrients, and

light. Ecology *89*: 2542-2552.

Vander Zanden, M. J., S. Chandra, S. Park, Y. Vadeboncoeur, and C. R. Goldman. 2006.

Efficiencies of benthic and pelagic trophic pathways in a subalpine lake. Canadian

Journal of Fisheries and Aquatic Sciences 63: 2608-2620.

Vanni, M. J. 2002. Nutrient cycling by animals in freshwater ecosystems. Annual Review of

Ecology and Systematics 33:341–370.

Vázquez, D. P., R. Ramos-Jiliberto, P. Urbani, and F. S. Valdovinos. 2015. A conceptual

framework for studying the strength of plant–animal mutualistic interactions. Ecology

Letters 18:385–400.

Welschmeyer, N. A. 1994. Fluorometric analysis of chlorophyll a in the presence of chlorophyll

b and pheopigments. Limnology and Oceanography 39:1985–1992.

Zipkin, E. F., P. J. Sullivan, E. G. Cooch, C. E. Kraft, B. J. Shuter, and B. C. Weidel. 2008.

Overcompensatory response of a smallmouth bass (Micropterus dolomieu) population to

harvest: release from competition? Canadian Journal of Fisheries and Aquatic Sciences

65:2279–2292.

**Figure 1** (a) Benthic gross primary production (estimated from $O_2$ production) in 55 experimental mesocosms (39 light and 16 dark) increases across the range of initial chironomid larval densities. Each point represents the oxygen production in a light-incubated mesocosm, measured on day 13 of the study. The solid line shows the linear relationship from the mixed model fitting the oxygen production of each mesocosm as a function of the initial number of chironomid larvae in a mesocosm. (b) Net oxygen change in the light-incubated mesocosms was used to estimate NPP in the mesocosms. These rates of change in oxygen concentrations were plotted against the number of larvae originally in each mesocosm. The solid line shows the linear fit of the mixed model predicting net oxygen change as a function of the number of chironomid larvae. (c) We obtained chlorophyll *a* concentrations from the top 0-1cm of sediments from the 55 experimental mesocosms. A linear model shows that chlorophyll *a* concentrations from experimental mesocosms increased significantly in response to higher numbers of chironomid larvae in the mesocosms (solid line). (d) We calculated the average dry mass of *C. islandicus* individuals (based on 30 haphazardly selected larvae) from each mesocosm containing chironomid larvae (n = 49 mesocosms). The dashed line shows the average initial dry weight of larvae before the start of the experiment (0.82 mg/individual). The solid line shows the fit of the linear model using the initial number of larvae as the predictor of *C. islandicus* final larval mass. (e) The rate of oxygen production per gram chlorophyll *a* in the light-incubated mesocosms increases as a function of the initial number of chironomid larvae in the mesocosm. The solid line shows the fit of the linear model using the number of chironomid larvae in each mesocosm as a predictor of oxygen production per gram of chlorophyll *a*.

**Figure 1:**
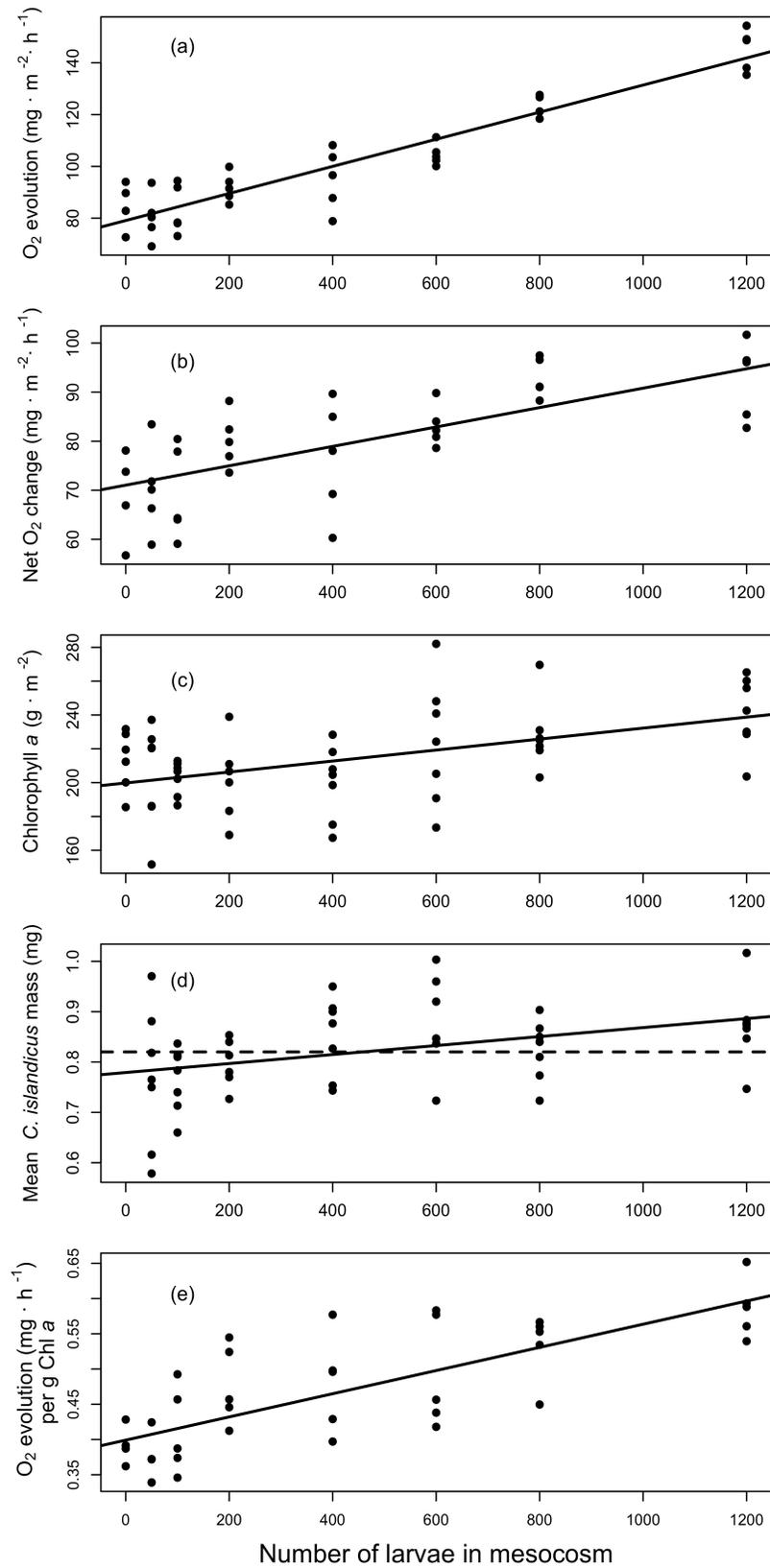
**Chapter 4:** Cohesion: A method for quantifying the connectivity of microbial communities

**Authors:** Cristina M. Herren[1], Katherine D. McMahon[2]


**Author affiliations:**

[1]Freshwater and Marine Sciences Program, University of Wisconsin - Madison, Madison, Wisconsin, USA

[2]Departments of Bacteriology and Civil and Environmental Engineering, University of Wisconsin - Madison, Madison, Wisconsin, USA

**Abstract**

The ability to predict microbial community dynamics lags behind the quantity of data available in these systems. Most predictive models use only environmental parameters, although a long history of ecological literature suggests that community complexity should also be an informative parameter. Thus, we hypothesize that incorporating information about a community's complexity might improve predictive power in microbial models. Here, we present a new metric, called community "cohesion," that quantifies the degree of connectivity of a microbial community. Here, we analyze six long-term (10+ year) microbial datasets using the cohesion metrics and validate our approach using datasets where absolute abundances of taxa are available. As a case study of our metrics' utility, we show that community cohesion is a strong predictor of Bray-Curtis dissimilarity ($R^2 = 0.47$) between phytoplankton communities in Lake Mendota, WI, USA. Our cohesion metrics outperform a model built using all available environmental data collected during a long-term sampling program. The result that cohesion corresponds strongly to Bray-Curtis dissimilarity is consistent across the six long-term time series, including five phytoplankton datasets and one bacterial 16S rRNA gene sequencing dataset. We explain here the calculation of our cohesion metrics and their potential uses in microbial ecology.

**Introduction**

Most efforts to model microbial communities primarily use environmental drivers as predictors of community dynamics (Patterson 2009, Hambright et al. 2015). However, despite the vast quantities of data becoming available about microbial communities, predictive power in microbial models is often surprisingly poor (Blaser et al. 2016). Even in one of most well-studied microbial systems, the San Pedro Ocean Time Series (SPOT), there are sampling sites where none of the 33 environmental variables measured are highly significant ($P < 0.01$) predictors of community similarity (Cram et al. 2015). Thus, there may be room to improve predictive models by adding new parameters; ecological literature has long suggested that the degree of complexity in a community should inform community dynamics (MacArthur 1955, Cohen and Newman. 1985, Wootton and Stouffer 2016). We use the term "complexity" as defined in the theoretical ecology literature, which refers to the number and strength of connections in a food web (May 1974). We hypothesize that incorporating information about the complexity of microbial communities could improve predictive power in these communities.

Here, we present a workflow to generate metrics quantifying the connectivity of a microbial community, which we call "cohesion." We demonstrate how our cohesion metrics can be used to predict community dynamics by showing that cohesion is significantly related to the rate of compositional turnover (Bray-Curtis dissimilarity) in microbial communities. As an application of our metrics, we present a case study using our newly developed cohesion variables as predictors of the compositional turnover rate (a common response variable in microbial ecology) in phytoplankton communities. Prior modeling efforts have indicated that incorporating taxon traits and interactions improved models of phytoplankton community assembly (Litchman and Klausmeier 2008, Thomas et al. 2012). However, even basic traits such as taxonomy are still

often unknown for other microbial taxa, such as bacteria (Newton et al. 2011). Thus, taxon interactions and community connectivity must be inferred statistically.

Our cohesion metrics overcome two barriers that often preclude using information about community complexity in microbial analyses. First, the large number of taxa in microbial datasets makes it difficult to use information about all taxa in statistical analyses. Although methods exist to analyze microbial community interconnectedness (e.g. Local Similarity Analysis, artificial neural networks), this often involves constructing networks with many parameters that are difficult to interpret. Second, microbial community data are often "relativized" or "compositional" datasets, where the abundance of each taxon represents the fraction of the community it comprises. This creates several problems in downstream analysis (Weiss et al. 2016). For example, taxon correlation values are different in absolute versus relative datasets (Faust and Raes 2012, Friedman and Alm 2012), and it is unclear how using relative abundances influences the apparent population dynamics of individual taxa (Lovell et al. 2015). Thus, these two features (many taxa and relative abundance) have previously proven problematic when analyzing microbial community connectivity. The methods used to account for these biases influence the results of the analyses. For instance, the proportion of positive versus negative pairwise interactions identified in a single dataset varied widely when using different correlation detection methods (Weiss et al. 2016). Additionally, the power to detect significant relationships between taxa declines steeply when taxa are less persistent and as relationships become non-linear (Weiss et al. 2016). In contrast to existing correlation detection methods, which aim to identify significant pairwise interactions, our cohesion metrics evaluate connectivity at the community level.

Here, we describe and test a method to quantify one aspect of microbial community complexity. Our resulting "cohesion" metrics quantify the connectivity of each sampled community. Thus, our cohesion metrics integrate easily with other statistical analyses and can be used by any microbial ecologist interested in asking whether community interconnectedness is important in their study system. We demonstrate how to obtain these cohesion metrics from time series data and, as a case study, show how cohesion relates to rates of compositional turnover in long-term microbial datasets. We develop this workflow with datasets where raw abundance data are available and use these raw abundances to validate our methods when working with relativized datasets. Thus, our approach was designed to overcome known challenges of analyzing microbial datasets.

## Methods and Results

### Description of datasets

The North Temperate Lakes Long Term Ecological Research (NTL LTER) database hosts many long-term time ecological series. We used five long-term phytoplankton datasets (two from the NTL LTER and three from the Cascade research group) to validate the cohesion workflow. These datasets met a number of criteria that made them good candidates for the validation: samples were collected regularly, sampling spanned multiple years and many environmental gradients, and taxa were counted in absolute abundance. The term "phytoplankton" refers to the polyphyletic assemblage of photosynthetic aquatic microbes (Litchman and Klausmeier 2008). The datasets are from the following lakes in Wisconsin, USA: Lake Mendota (293 samples with 410 taxa over 19 years), Lake Monona (264 samples with 382 taxa over 19 years), Paul Lake (197 samples with 209 taxa over 12 years), Peter Lake (197

samples with 237 taxa over 12 years), and Tuesday Lake (115 samples with 121 taxa over 12 years). These lakes vary in size, productivity, and food web structure. Lake Mendota and Lake Monona are large (39.4km$^2$ and 13.8km$^2$), urban, eutrophic lakes (Brock 2012). Peter, Paul, and Tuesday lakes are small (each < 0.03km$^2$) lakes surrounded by forest (Carpenter and Kitchell 1996). Peter Lake and Tuesday Lake were also subjected to whole-lake food web manipulations during the sampling timeframe (detailed in Elser and Carpenter 1988 and Cottingham et al. 1998). After validating our workflow using the phytoplankton datasets, we tested the cohesion metrics on a bacterial dataset obtained using 16S rRNA gene amplicon sequencing. These types of datasets often contain thousands of taxa, most of them rare, which may influence the results of correlation-based analyses (Faust and Raes 2012). We used the Lake Mendota bacterial 16S rRNA gene sequencing time series (91 samples with 7081 taxa over 11 years) for this analysis (Hall et al in review). Sample processing, sequencing and core amplicon data analysis were performed by the Earth Microbiome Project (EMP) (www.earthmicrobiome.org) (Gilbert et al 2014), and all amplicon sequence data and metadata have been made public through the data portal (qiita.microbio.me/emp). Briefly, community DNA (Kara et al 2013) was used to amplify partial 16S rRNA genes using the 515F-806R primer pair (Caporaso et al 2011) and an Illumina MiSeq, with standard EMP protocols.

We present the workflow using results from the Lake Mendota phytoplankton dataset, as it is the largest (longest duration and most taxa) dataset available in absolute abundance. The dominant taxa in the Lake Mendota phytoplankton dataset change throughout the year, with diatoms most abundant during the spring bloom and cyanobacteria most abundant in summer. Details about phytoplankton datasets can be found at https://lter.limnology.wisc.edu/. Further details about the Lake Mendota 16S rRNA gene dataset are included in the SOM.

*Data curation*

Phytoplankton densities in Lake Mendota varied by more than 2 orders of magnitude

between sample dates. Densities of cells in these samples ranged from 956 cells/mL to 272 281

cells/mL. We removed individuals that were not identified at any level (e.g. categorized as

Miscellaneous). For each sample date, we converted the raw abundances to relative abundances

by dividing each taxon abundance by the total number of individuals in the community, such that

all rows summed to 1. Relative abundances indicate the fraction of a community comprised by

the taxon. We removed taxa that were not present in at least 5% of samples, as we were not

confident that we could recover robust connectedness estimates for very rare taxa. This cutoff

retained an average of 98.9% of the identified cells in each sample. Results of our analyses using

other cutoff values can be found in the supplementary online material (SOM).

*Overview*

The input of our workflow is the taxon relative abundance table, and the outputs are

measurements of the connectivity of each sampled community, which we call community

"cohesion" (Fig. 1). In the process, our workflow also produces metrics of the connectedness of

each taxon. Briefly, our workflow begins by calculating the pairwise correlation matrix between

taxa, using all samples. We use a null model to account for bias in these correlations due to the

skewed distribution of taxon abundances (i.e. many small values and a few large values) and

relativized nature of the dataset (i.e. all rows sum to 1). We subtract off these "expected"

correlations generated from the null model to obtain a matrix of corrected correlations. For each

taxon, the average positive corrected correlation and average negative corrected correlation are

recorded as the connectedness values. As previously noted, our goal was to create a metric of

connectivity for each community; thus, the next step in the workflow calculates cohesion values

for each sample. Cohesion is calculated by multiplying the abundance of each taxon in a sample by its associated connectedness values, then summing the products of all taxa in a sample. There are two metrics of cohesion, because we separately calculate metrics based on the positive and negative relationships between taxa. Within each section (1, 2, and 3), we alternate between presenting an analysis step and showing a validation of these techniques.

*1. CONNECTEDNESS METRIC*

Analysis

*Null Models*

It is difficult to directly observe interactions within microbial communities, so correlations are often used to infer relationships between taxa or between a taxon and the environment. Thus, we used a correlation-based approach for determining the connectedness of taxa. However, when using correlation-based approaches with relativized microbial datasets, it is necessary to use a null model to evaluate how the features of the dataset (skewed abundances and the fact that all rows sum to 1) contribute to correlations between taxa (Weiss et al. 2016). The purpose of a null model is to assess the expected strengths of correlations when there are no true relationships between taxa (Ulrich and Gotelli 2010).

The null model was used to calculate how strongly the features common to microbial datasets contribute to taxon connectedness estimates, so that this structural effect could be subtracted from the connectedness metrics. Of the several dozen null models tested, we have selected two for inclusion in the cohesion R script. We discuss both null models here and in the SOM. The SOM and readme document should assist in choosing the null model appropriate for a given dataset. While testing various null models, it became clear that a taxon's pairwise

correlation values were strongly related to its persistence (fraction of samples when present) across the dataset. Thus, taxon persistence was preserved in both null models.

The objective of the null model was to calculate the strength of pairwise correlations that would be observed if there were no true relationship between taxa. This paragraph describes the "taxon/column shuffling" null model used for the phytoplankton dataset analyses. During each iteration, one taxon was designated as the "focal taxon" (Fig. 2). For each taxon *besides the focal taxon*, abundances in the null matrix were permuted (i.e. randomly sampled without replacement) from their abundance distribution across all samples. Then, we calculated Pearson correlations between the focal taxon and the randomized other taxa. We iterated through this process of calculating pairwise correlations between the focal taxon and all other taxa 200 times. The median correlations from these 200 randomizations were called the "expected" correlations for the focal taxon. We recorded the median value as the "expected" correlation, rather than the mean value, because distributions were skewed toward larger values. Thus, a greater proportion of the distribution fell within one standard deviation of the median, as compared to within one standard deviation of the mean. We repeated this process for each taxon as the focal taxon, which resulted in a matrix of expected taxon correlations. Finally, we subtracted the expected taxon correlations from their paired observed taxon correlations, thereby producing a matrix where each value was an observed minus expected correlation for the given pair of taxa.

The second null model uses the same workflow as described above, where the dataset is iteratively randomized and median correlations are used as the "expected" pairwise correlations. However, the method of randomization is different; instead, the abundances of all taxa present within one *sample* were randomized. We refer to this null model as the "row shuffling" model. The benefit of this null model is that row sums are maintained. Thus, negative dependencies

between taxa within the same sample are accounted for in this model. The drawback of this null model is that a taxon might be assigned an abundance value that is implausible (i.e. larger than its maximum observed abundance). In the online script to calculate cohesion, we have included the option to choose between these two null models (taxon shuffle and row shuffle).

We have included an additional option to input a pre-determined correlation matrix, thereby bypassing the null model. Using a pre-determined correlation matrix allows researchers to use a different correlation detection strategy to generate the correlation matrix. This option to import a custom correlation matrix makes our cohesion workflow compatible with other software packages designed for detecting pairwise relationships in microbial communities.

*Calculating Connectedness*

We calculated taxon connectedness values from the corrected (observed minus expected) correlation matrix. For each taxon, we separately averaged its positive and negative correlations with other taxa to produce a value of positive connectedness and a value of negative connectedness. We kept positive and negative values separate for both mathematical and biological reasons. First, we had hypothesized that positive and negative correlations may capture different ecological relationships between taxa. Furthermore, positive correlations were stronger (an average of 2.5 times larger in magnitude) than negative correlations. And, correlation distributions were generally skewed toward positive values. Thus, a small number of positive correlations could mute the signal of negative correlations, if positive and negative correlations were averaged together.

The averaging step in this workflow was intended to overcome the issue that individual correlations between taxa can be influenced by many factors and may be spurious (Fisher and Mehta 2014). However, assuming that correlations often (but not always) reflect complexity in a

community, the average of many correlations should be a more robust metric of complexity than any single correlation. In other words, we assume only that highly connected taxa have stronger correlations *on average*. Invoking the law of large numbers, these average correlations should be increasingly accurate measures of a taxon's connectedness as the number of pairwise correlations increases (i.e. as the number of taxa in the dataset increases). Similarly, applying the Central Limit Theorem, each mean correlation should be normally distributed with low variance due to the large number of pairwise correlations.

Validation

As discussed previously, there are inherent limitations of using correlation-based methods with relative abundance data instead of absolute counts (Fisher and Mehta 2014). Thus, we examined whether a measure of connectedness based on absolute abundance would show the same pattern observed using the relativized data. However, we needed a different approach for calculating correlations in order to account for the following properties of count data: 1) variance-mean scaling, which results in very large population variances of abundant taxa (Taylor 1961) and 2) the fact that individual population sizes are strongly related to overall community densities, which causes positive correlations among all taxa (Doak et al. 1998). As noted previously, phytoplankton densities in Lake Mendota samples varied by more than 2 orders of magnitude among sample dates. Therefore, using correlations between raw abundances would inflate the positive relationships between taxa as a result of changing overall community density. Thus, we first detrended the count data to account for changing community density (on different sampling dates) and drastically different variances of taxon populations (which are expected as a result of mean-variance scaling).

We used a hierarchical linear model to estimate the effects of overall community density and mean taxon abundance on individual taxon observations (sensu Jackson et al. 2012), so that these effects could be removed when calculating correlations. We modeled the abundance of each taxon at each time point as a function of sample date and taxon, assuming a quasipoisson distribution (which accounts for increases in population variances when population means increase). The model estimates a mean abundance effect for each sample, based on the abundances of each taxon in the sample. Similarly, the model estimates mean abundances for each taxon, based on the distribution of taxon abundances across all samples. Thus, the residuals of this analysis represent the normalized (transformed) deviations of taxon abundances after accounting for overall community density on the sample date and taxon abundance/variance. We created a pairwise correlation matrix for the phytoplankton taxa using the correlations between these residuals. We calculated connectedness metrics from the pairwise correlation matrix using the same technique that we applied to the corrected correlation matrix from the relativized data: we used the average positive and negative taxon correlations as their connectedness values.

We validated our workflow for the relative abundance dataset using the estimates of taxon connectedness obtained from the absolute abundance dataset. Comparing the connectedness values from these two methods shows strong agreement between the two sets of connectedness metrics (correlation for positive connectedness metrics = 0.820; correlation for negative connectedness metrics = 0.741, Fig. 3). Although two taxa deviate from the linear relationship between the negative connectedness metrics (appearing as outliers in Fig. 3B), both metrics rate these taxa as having strong connectedness arising from negative correlations. Thus, the two methods are qualitatively consistent for these two anomalous points. Furthermore, using the null model improved the correspondence between absolute and relative connectedness

metrics, as measured by their proportionality. The variance in the proportions (relative metric / absolute metric) decreased after the null model correction was implemented (variance in proportions for positive metrics: uncorrected = 0.25, corrected = 0.065; variance in proportions for negative metrics: uncorrected = 0.047, corrected = 0.035).

## 2. COHESION METRIC

Analysis

Many researchers seek to detect differences in community connectivity across time, space, or treatments. Thus, it would be useful to have a metric that quantifies, for each community, the degree to which its component taxa are connected. The aim of our cohesion metric is to quantify the instantaneous connectivity of a community, where connectivity increases when highly connected taxa become more abundant in the community. We used a simple algorithm to collapse the connectedness values of individual taxa into two parameters representing the connectivity of the entire sampled community, termed "cohesion." Again, one metric of cohesion stems from positive correlations, and one metric stems from negative correlations. To calculate each cohesion metric, we multiplied the relative abundance of taxa in a sample by their associated connectedness values and summed these products. This cohesion index can be represented mathematically as the sum of the contribution of each of the $n$ taxa in the community, after removing rare taxa (Eq. 1). Thus, communities with high relative abundances of strongly connected taxa would have a high score of community cohesion. We note that this index is bounded by -1 to 0 for negative cohesion or from 0 to 1 for positive cohesion.

$$cohesion = \sum_{i=1}^{n} abundance_i * connectedness_i \qquad \text{Eq. 1}$$

<u>Validation</u>

We had hypothesized that our cohesion metrics could be significant predictors of microbial community dynamics. Thus, a natural question to ask was whether our metrics of cohesion outperform environmental variables when analyzing the Lake Mendota phytoplankton data. Fortunately, the NTL LTER program has collected paired environmental data for the Lake Mendota phytoplankton samples. We obtained these environmental datasets to use as alternative predictors of phytoplankton community dynamics in Lake Mendota. The environmental datasets available (11 variables) were: water temperature, air temperature, dissolved oxygen concentration, dissolved oxygen saturation, Secchi depth, combined $NO_3 + NO_2$ concentrations, $NH_4$ concentration, total nitrogen concentration, dissolved reactive phosphorus concentration, total phosphorus concentration, and dissolved silica concentrations. Protocols, data, and associated metadata can be found at https://lter.limnology.wisc.edu/. We use these environmental data to build an alternate model in our case study below.

*3. CASE STUDY OF UTILITY*

<u>Analysis</u>

To demonstrate their utility, we applied our new metrics to the Lake Mendota phytoplankton dataset. We tested whether community cohesion could predict compositional turnover, a common response variable in microbial ecology. We used multiple regression to model compositional turnover (Bray-Curtis dissimilarity between time points) as a function of community cohesion at the initial time point. That is, Bray-Curis dissimilarity was the dependent

variable, while positive and negative cohesion were the independent variables. Because time

between samples influences Bray-Curtis dissimilarity (Nekola and White 1999, Shade et al.

2013), we only included pairs of samples taken within 36 to 48 days of each other. These criteria

included 186 paired communities across the 19 years. Cohesion values (both positive and

negative) were calculated at the first time point for each sample pair. We chose this timeframe

because it was sufficiently long for multiple phytoplankton generations to have occurred, and

because this timeframe was compatible with the sampling frequency.

Community cohesion was a strong predictor of compositional turnover (Fig. 4). The

regression using our cohesion metrics explained 46.5% of variability (adjusted $R^2 = 0.465$) in

Bray-Curtis dissimilarity. Cohesion arising from negative correlations was a highly significant

predictor, whereas cohesion arising from positive correlations was not significant (negative

cohesion: $F_{1, 183} = 6.81$, $p < 1* 10^{-20}$; positive cohesion: $F_{1, 183} = 0.735$, $p = 0.405$).

For the purpose of model comparison, we used the associated environmental data to

model Bray-Curtis dissimilarity as a function of environmental drivers. We included as

predictors the 11 variables previously mentioned, as well as 11 additional predictors that

measured the change in each of these variables between the two sample dates. Finally, because

many chemical and biological processes are dependent on temperature (Brown et al. 2004), we

included first order interactions between water temperature and the 21 other variables. We first

included all 43 terms in the model, then used backward selection (which iteratively removes the

least significant term in the model, beginning with interaction terms) until all remaining terms in

the model were significant at $p < 0.1$, as to maximize the adjusted $R^2$ value. Although this

analysis does not represent an exhaustive list of possible environmental drivers, it includes all

available paired environmental data from the long-term monitoring program. Twenty-nine values

of Bray-Curtis dissimilarity were excluded from this analysis (leaving 157 of the 186 values), because they lacked one or more associated environmental variables. Additional details about this analysis can be found in the SOM.

In the final model after backward selection, 16 variables were retained as significant predictors (see SOM). Significance was determined using type III sums of squares. Using the guideline that each variable should have approximately 10 additional data points to prevent overparameterization (Peduzzi et al. 1995), we were not concerned about overfitting. The adjusted $R^2$ of this model was 0.229. The *non-adjusted* $R^2$ value of the full model (all 43 variables) was 0.393. When adding negative cohesion as a parameter into the final environmental model, negative cohesion was highly significant ($p < 1*10^{-13}$) and 12 environmental variables remained significant at $p < 0.1$.

To address the generality of the relationship between cohesion and community turnover, we calculated cohesion metrics and Bray-Curtis dissimilarity for the four other phytoplankton datasets (Monona, Peter, Paul, and Tuesday lakes) and for the Lake Mendota bacterial 16S rRNA gene sequencing dataset. Community cohesion was a significant predictor of Bray-Curtis dissimilarity in all datasets. In each instance, stronger cohesion resulting from negative correlations was related to lower compositional turnover. Table 1 presents the results of these analyses and associated workflow parameters. Additional information about the sensitivity of model performance to varying parameters can be found in the SOM.

Validation

Strong correlations between predictor variables are known to influence the results of statistical analyses (Neter et al. 1996). Thus, we wondered whether strong correlations between taxa would necessarily generate the observed relationship where greater cohesion is related to

lower compositional turnover. We conducted simulation studies to investigate whether our significant results might be simply an artifact of strong inter-taxon correlations. We generated datasets where taxa were highly correlated in abundance, as if they were synchronously responding to exogenous forces. We calculated cohesion metrics and Bray-Curtis dissimilarities for the simulated datasets to analyze whether strong taxon correlations was sufficient to produce results similar to those we observed in the real data.

Here, we briefly describe the process used to simulate datasets, while additional details can be found in SOM. First, we generated four autocorrelated vectors to represent exogenous forces, such as environmental drivers. Taxa were artificially correlated to these external vectors, thereby also producing strong correlations between taxa. We manipulated the taxon abundances to mimic other important features of the microbial datasets, including skewed taxon mean abundances and a large proportion of zeroes in the dataset. We calculated cohesion metrics and Bray-Curtis dissimilarities for the simulated datasets, and we used a multiple regression to model Bray-Curtis dissimilarity as a function of positive cohesion and negative cohesion. We recorded the $R^2$ value and parameter estimates of this multiple regression. We repeated this simulation process 500 times to generate distributions of these results.

Our cohesion metrics had a very low ability to explain compositional turnover (Bray-Curtis dissimilarity) in the simulated datasets. The median model adjusted $R^2$ value was 0.022, with 95% of adjusted $R^2$ values below 0.088 (Fig. 5). Although the community cohesion metrics were highly significant predictors ($p < 0.001$) of community turnover more commonly than would be expected by chance (1.0% of simulations for positive cohesion and 8.6% for negative cohesion), the proportion of variance explained by these metrics was comparatively very low. For comparison, across the six long-term datasets from Wisconsin lakes, model adjusted $R^2$

values ranged from 0.36 to 0.50. Thus, there was comparatively little ability to explain compositional turnover in the simulated datasets using our cohesion metrics.

**Discussion**

The ability to predict microbial community dynamics lags behind the amount of data collected in these systems (Blaser et al. 2016). Here, we present new metrics, called "cohesion," which can be used as additional predictor variables in microbial models. The cohesion metrics contain information about the connectivity of microbial communities, which has been previously hypothesized to influence community dynamics (MacArthur 1955, May 1972, Nilsson and McCann 2016). Our cohesion metrics are easily calculated from a relative abundance table (R script provided online) and might be of interest to a variety of microbial ecologists and modelers.

In the Lake Mendota phytoplankton example, our two cohesion parameters outperformed the available environmental data at predicting phytoplankton community changes. The two cohesion parameters explained 46.5% of variability (adjusted $R^2 = 0.465$) in community turnover over 19 years of phytoplankton sampling, in comparison to the final environmental model using 16 predictors, which explained 22.9% of community turnover (adjusted $R^2 = 0.229$). The simultaneous significance of negative cohesion and 12 environmental variables when all predictors were included in a single model indicates that environmental variables and negative cohesion explained different sources of variability in Bray-Curtis dissimilarity. Although there are almost certainly important predictors missing from the environmental model (e.g. photosynthetically active radiation, three-way interactions), the environmental model represents a commonly applied approach to explaining microbial compositional turnover (Tripathi et al. 2012, Chow et al. 2013) that uses all associated environmental data from a long-term sampling

program. Although we still strongly advocate for the collection of environmental data, we note

that cohesion was a much better predictor of compositional turnover than any available

environmental variable.

Our workflow overcomes many challenges associated with using correlation-based

techniques in microbial datasets. The validations we conducted indicated that our connectedness

metrics are appropriate for relativized datasets, because connectedness metrics from relative and

absolute datasets showed strong correspondence. Most DNA sequencing datasets are only

available in relative abundance. Previous methods for analyzing relative abundance datasets have

identified potential pitfalls of calculating correlations for these data (Friedman and Alm 2012,

Weiss et al. 2016); however, the extent to which these biases influence analysis results is often

unknown, because paired absolute abundance datasets do not exist. The validations of our

cohesion workflow with absolute abundance data indicate that the steps taken to account for

biases (using a null model and averaging pairwise correlations) make the cohesion metrics robust

for relative abundance datasets.

Our cohesion metrics address a common problem of techniques describing community

complexity (such as network analyses), which is that they do not quantify the connectivity of

individual communities. For instance, the "hairball" generated from a network analysis is

generated from many samples; there are no parameters specific to each sample, and therefore the

network cannot be used as a predictor variable. Thus, existing methods to quantify connectivity

do not pair easily with other analyses. Furthermore, in contrast to many other network analyses,

we did not attempt to calculate significance values for pairwise correlations as a part of the

cohesion workflow. Based on our *a priori* hypothesis that weak interactions are ecologically

important (McCann et al. 1998), we included all pairwise correlations in the connectedness

metrics. Our cohesion metrics quantify sample connectivity using only two parameters, which can be used as predictors in a variety of further analyses (linear regression, ordinations, time series, etc.). Finally, our simulations showed that strong inter-taxon correlations were not sufficient to reproduce the observed result that cohesion was a strong predictor of Bray-Curtis dissimilarity. In the simulations, cohesion had low explanatory power, even though taxa were highly correlated. From this result, we infer that correlations between taxa in real communities are an important aspect of complexity that is captured by our cohesion metrics.

Our cohesion metrics explain a significant amount of compositional change in all six datasets (five phytoplankton and one bacterial 16S rRNA gene dataset). Yet, it is not immediately clear what cohesion is measuring. There are two broad factors that could cause correlations between taxa: biotic interactions and environmental drivers. Thus, at least one of these two factors must underlie our connectedness and cohesion metrics. Here, we discuss the evidence supporting either of these interpretations:

*Cohesion as a Measure of Biotic Interactions*

Even if shared responses to environmental drivers underlie most pairwise taxon correlations, cohesion could still indicate biotic interaction strength in a community. This would occur if taxa were influenced to the same degree by environmental drivers, but differentially influenced by species interactions. In this case, averaging over all correlations would give larger connectedness values for strong interactors and smaller connectedness values for weak interactors. Many studies have indicated that microbial taxa have differential interaction strengths. For example, some microbial communities contain keystone taxa, which have disproportionate effects on community dynamics through their strong taxon interactions (Trosvik and de Muinck 2015, Banerjee et al. 2016). Similarly, recent work suggests that many taxa

within candidate phyla are obligate symbionts, meaning they must interact strongly with other taxa for their survival and reproduction (Kantor et al. 2013, Hug et al. 2016). Conversely, there are many taxa that can be modeled adequately as a function of environmental drivers; this is true for some bloom forming cyanobacteria, which are known to respond strongly to nutrient concentrations and temperature (McQueen and Lean 1987, Beaulieu et al. 2013). Taken together, these studies suggest that there is a wide spectrum of how strongly taxa interact with one another. These differences in interaction strength would be detected by our connectedness metric due to averaging over the large number of pairwise correlations. Thus, it is plausible that connectedness and cohesion are reflecting biotic interactions in communities.

We now examine results from the long-term dataset analyses under the assumption that cohesion measures biotic interactions. The Bray-Curtis dissimilarity regression results would mean that communities with many strong interactors have lower rates of change, especially when the interactions create negative correlations between taxon abundances. This finding is in line with prior work showing that biotic interactions affect microbial community stability (Coyte et al. 2016). Thus, the interpretation that stronger biotic interactions lead to lower compositional turnover is a plausible explanation for our observed results. However, we specifically refrain from interpreting positive or negative connectedness values as indications of specific biotic interactions, such as predation, competition, or mutualism. For example, a positive correlation between two taxa could be the result of a mutualism between the taxa, or it could be the result of a shared predator declining in abundance. Further work, both empirical and theoretical, is necessary to identify what these positive and negative correlations signify in the context of the ecology of these organisms.

*Cohesion as a Measure of Environmental Synchrony*

We now consider the possibility that connectedness and cohesion are simply detecting environmental synchrony. If a subset of taxa respond to a changing environmental driver, then these taxa will have strong pairwise correlations. For example, correlations between phytoplankton species of the same genus (and, therefore, with similar niches) can be upwards of 0.9, indicating strong similarity in abundance patterns. In this case, connectedness would measure the degree of environmentally-driven population synchrony that a taxon has with other taxa. A high cohesion value would indicate that a community has many taxa that respond simultaneously to external forces; then, cohesion would quantify overall community responsiveness to one or more environmental drivers. Under this assumption, cohesion should correlate with environmental drivers (e.g. cohesion is high because many taxa are positively correlated to warm temperatures, but cohesion drops when it gets colder and these taxa senesce). We tested this prediction with 22 variables from the environmental model (11 for the environmental variables and 11 for the changes in environmental variables) and found that negative cohesion in the Lake Mendota phytoplankton dataset generally had weak correlations with these predictors (absolute correlations < 0.25, SOM). We also looked for a seasonal trend in cohesion, but found no significant correlation between cohesion (positive or negative) and Julian Day, or a quadratic term for Julian Day. Thus, we do not find any evidence that cohesion is simply reproducing the information contained in environmental data. Finally, our simulations show one example where taxon abundances could be driven exclusively by external factors (such as the environment), but this does not necessarily lead to strong predictive power of compositional turnover. However, our simulations omitted many features of real ecological

communities, and so we cannot completely rule out the possibility that environmental drivers contributed to our cohesion metrics in the phytoplankton datasets.

Under the assumption that cohesion measures environmentally driven population synchrony, we examine our result that stronger negative cohesion was related to lower Bray-Curtis dissimilarity. In this scenario, communities that have strong cohesion contain high abundances of taxa that respond simultaneously to environmental forces. Then, communities with many synchronous taxa would turn over more slowly than communities with taxa whose abundances are independent of the environment. This conclusion is counterintuitive, but possible. This pattern could occur if taxa that are strongly influenced by the environment have lower variability than taxa that are weakly influenced by the environment; in that case, highly correlated taxa would have their abundances more tightly regulated than other taxa. Although possible, this explanation disagrees with many studies that have found that environmental gradients regulate which taxa can persist in communities (Fierer and Jackson 2006, Walter and Ley 2011, Freedman and Zak 2015).

Comparing the two possible signals that cohesion might be detecting, we believe the evidence points to biotic interaction as the larger contributor. However, we expect that environmental synchrony is captured to some extent, with the relative importance of environmental factors depending on the particular communities and ecosystem. In instances where synchronous responses to environmental drivers cause positive correlations between taxa, we would expect this environmentally-driven signal to affect positive cohesion values more than negative cohesion values. Regardless of the ecological force measured by cohesion, there is a clear result in the six datasets analyzed that stronger negative cohesion is related to lower compositional turnover. This result suggests that negative correlations between taxa are arranged

non-randomly to counteract one another, thereby stabilizing community composition. In other words, relationships between taxa appear to buffer, rather than amplify, changes to community composition. This result agrees with prior theoretical models that propose that feedback loops originating from taxon interactions are integral to modulating food web stability (Neutel et al. 2007, Brose 2008). Although stronger negative cohesion was related to lower compositional turnover, negative pairwise correlations were, on average, weak. The negative connectedness values ranged from -0.004 to -0.12, and the mean negative correlation was -0.022. Thus, our results are not inconsistent with the hypothesis that weak interactions are stabilizing to communities (McCann et al. 1998). The finding that negative cohesion was stabilizing was not easily replicated in our simulations, where positive and negative correlations were interspersed with random magnitude throughout the dataset. Thus, the arrangement of correlations between taxa in the dataset appears to be an important feature of real communities that may contribute to their stability (Worm and Duffy 2003).

*Guidelines for Using Our Metrics*

Although we used long-term time series datasets for the analyses presented here, our cohesion metrics can be used to predict community dynamics in a variety of datasets. For example, cohesion could be used with a spatially explicit dataset, where samples were collected from different locations across a landscape. In the context of phytoplankton samples, this could be a dataset consisting of samples from different locations in a lake or watershed. Then, the cohesion metrics could be used to predict community composition change at one location over time, or to predict differences in community composition between locations. It would also be interesting to investigate how cohesion is affected by experimental perturbations. Finally, cohesion could be used as a predictor in of many response variables. Additional applications of

the cohesion metrics could include identifying communities susceptible to major compositional change (e.g. cyanobacterial blooms, infection in the human microbiome), relating community cohesion to spatial structure (e.g. how taxon connectedness relates to the dispersal abilities of different microbial taxa), and investigating how disturbance influences cohesion (e.g. how illness influences the cohesion of communities in a host-associated microbiome, how oil spills affect cohesion of marine microbial communities). The consistent results between the phytoplankton datasets and the bacterial 16S rRNA gene dataset indicates that our cohesion metrics are robust for DNA sequencing datasets.

The critical step in the cohesion workflow is calculating reliable correlations between taxa. Thus, some datasets will be more suitable for our cohesion metric than others. For example, a dataset consisting of 20 samples from five lakes over multiple years might be a poor candidate for the cohesion metrics. In this case, correlations between taxa might be driven mainly by environmental differences or location, and the sample number would be too low to calculate robust correlations. Based on the phytoplankton datasets analyzed here, we suggest a lower limit of 40-50 samples when calculating cohesion metrics, with more samples necessary with more heterogeneous datasets. We also suggest including environmental variables as covariates when analyzing heterogeneous datasets. Finally, the persistence cutoff for including taxa should be adjusted based on the dataset being analyzed. For example, in datasets obtained by DNA sequencing, the sequencing depth affects taxon persistence (Smith and Peay 2014). Thus, for DNA sequencing datasets, we also recommend implementing a cutoff by mean abundance, where very rare taxa are omitted from the cohesion metrics.

**Conclusion**

Our cohesion metrics provide a method to incorporate information about microbial community complexity into predictive models. These metrics are easy to calculate, needing only a relative abundance table. Furthermore, across all datasets analyzed in this study, negative cohesion was strongly related to compositional turnover. In systems where cohesion is a significant predictor of community properties (e.g. nutrient flux, rates of photosynthesis), this result could guide further investigation into the effects of microbial interactions in mediating community function. In this case, researchers might focus their efforts on understanding the role of highly connected taxa, which are identified in our workflow. We aim to eventually determine the features that distinguish systems in which cohesion is important versus systems in which cohesion does not predict community properties.

**References**

Banerjee S, Kirkby CA, Schmutter D, Bissett A, Kirkegaard JA, Richardson AE. (2016).
Network analysis reveals functional redundancy and keystone taxa amongst bacterial and
fungal communities during organic matter decomposition in an arable soil. *Soil Biology
and Biochemistry* **97**: 188–198.

Beaulieu M, Pick F, Gregory-Eaves I. (2013). Nutrients and water temperature are significant
predictors of cyanobacterial biomass in a 1147 lakes data set. *Limnol Oceanogr* **58**:
1736–1746.

Blaser MJ, Cardon ZG, Cho MK, Dangl JL, Donohue TJ, Green JL, *et al*. (2016). Toward a
Predictive Understanding of Earth's Microbiomes to Address 21st Century Challenges.
*mBio* **7**: e00714–16.

Brock TD. (2012). A Eutrophic Lake: Lake Mendota, Wisconsin. Springer Science & Business
Media.

Brose U. (2008). Complex food webs prevent competitive exclusion among producer species.
*Proceedings of the Royal Society of London B: Biological Sciences* **275**: 2507–2514.

Brown JH, Gillooly JF, Allen AP, Savage VM, West GB. (2004). Toward a Metabolic Theory of
Ecology. *Ecology* **85**: 1771–1789.

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Lozupone CA, Turnbaugh PJ, *et al*.
(2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per
sample. *Proceedings of the National Academy of Sciences*, **108:** 4516-4522.

Carpenter SR, Kitchell JF. (1996). The Trophic Cascade in Lakes. Cambridge University Press.

Chow C-ET, Sachdeva R, Cram JA, Steele JA, Needham DM, 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.

Cohen JE, Newman CM. (1985). When will a large complex system be stable? *Journal of Theoretical Biology* **113**: 153–156.

Cottingham KL, Carpenter SR, Amand ALS. (1998). Responses of epilimnetic phytoplankton to experimental nutrient enrichment in three small seepage lakes. *J Plankton Res* **20**: 1889–1914.

Coyte KZ, Schluter J, Foster KR. (2015). The ecology of the microbiome: Networks, competition, and stability. *Science* **350**: 663–666.

Cram JA, Chow C-ET, Sachdeva R, Needham DM, Parada AE, Steele JA, *et al*. (2015). Seasonal and interannual variability of the marine bacterioplankton community throughout the water column over ten years. *ISME J* **9**: 563–580.

David Hambright K, 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.

Doak DF, Bigger D, Harding EK, Marvier MA, O'Malley RE, Thomson D. (1998). The Statistical Inevitability of Stability‐Diversity Relationships in Community Ecology. *The American Naturalist* **151**: 264–276.

Elser JJ, Carpenter SR. Predation-driven dynamics of zooplankton and phytoplankton communities in a whole-lake experiment. *Oecologia* **76**: 148–154.

Faust K, Raes J. (2012). Microbial interactions: from networks to models. *Nat Rev Micro* **10**: 538–550.

Fierer N, Jackson RB. (2006). The diversity and biogeography of soil bacterial communities. *PNAS* **103**: 626–631.

Fisher CK, Mehta P. (2014). Identifying Keystone Species in the Human Gut Microbiome from Metagenomic Timeseries Using Sparse Linear Regression. *PLOS ONE* **9**: e102451.

Freedman Z, Zak DR. (2015). Soil bacterial communities are shaped by temporal and environmental filtering: evidence from a long-term chronosequence. *Environ Microbiol* **17**: 3208–3218.

Friedman J, Alm EJ. (2012). Inferring Correlation Networks from Genomic Survey Data. *PLoS Computational Biology* **8**: 1–11.

Gilbert JA, Jansson JK, Knight R. (2014). The Earth Microbiome project: successes and aspirations. *BMC Biology* **12**: 69.

Hall MW, Rohwer RR, Perrie J, McMahon KD, Beiko RG. (in review) Ananke: Temporal clustering reveals ecological dynamics of microbial communities. Preprint available: https://peerj.com/preprints/2879/

Hug LA, Baker BJ, Anantharaman K, Brown CT, Probst AJ, Castelle CJ, *et al*. (2016). A new view of the tree of life. *Nature Microbiology* **1**: 16048.

Jackson MM, Turner MG, Pearson SM, Ives AR. (2012). Seeing the forest and the trees: multilevel models reveal both species and community patterns. *Ecosphere* **3**: 1–16.

Kantor RS, Wrighton KC, Handley KM, Sharon I, Hug LA, Castelle CJ, *et al*. (2013). Small Genomes and Sparse Metabolisms of Sediment-Associated Bacteria from Four Candidate Phyla. *mBio* **4**: e00708–13.

Kara EL, Hanson PC, Hu YH, Winslow L, McMahon KD. (2013). A decade of seasonal dynamics and co-occurrences within freshwater bacterioplankton communities from eutrophic Lake Mendota, WI, USA. *The ISME journal*, **7(3):** 680-684.

Litchman E, Klausmeier CA. (2008). Trait-Based Community Ecology of Phytoplankton. *Annual Review of Ecology, Evolution, and Systematics* **39**: 615–639.

Lovell D, Pawlowsky-Glahn V, Egozcue JJ, Marguerat S, Bähler J. (2015). Proportionality: A Valid Alternative to Correlation for Relative Data. *PLOS Comput Biol* **11**: e1004075.

MacArthur R. (1955). Fluctuations of Animal Populations and a Measure of Community Stability. *Ecology* **36**: 533–536.

May RM. (1972). Will a Large Complex System be Stable? *Nature* **238**: 413–414.

May RM. (1974). Stability and Complexity in Model Ecosystems. Princeton University Press.

McCann K, Hastings A, Huxel GR. (1998). Weak trophic interactions and the balance of nature. *Nature* **395**: 794–798.

McQueen DJ, Lean DRS. (1987). Influence of Water Temperature and Nitrogen to Phosphorus Ratios on the Dominance of Blue-Green Algae in Lake St. George, Ontario. *Can J Fish Aquat Sci* **44**: 598–604.

Nekola JC, White PS. (1999). The distance decay of similarity in biogeography and ecology. *Journal of Biogeography* **26**: 867–878.

Neter J, Kutner MH, Nachtsheim CJ, Wasserman W. (1996). *Applied linear statistical models* (Vol. 4). Chicago: Irwin.

Neutel A-M, Heesterbeek JAP, van de Koppel J, Hoenderboom G, Vos A, Kaldeway C, *et al*. (2007). Reconciling complexity with stability in naturally assembling food webs. *Nature* **449**: 599–602.

Newton RJ, Jones SE, Eiler A, McMahon KD, Bertilsson S. (2011). A Guide to the Natural History of Freshwater Lake Bacteria. *Microbiol Mol Biol Rev* **75**: 14–49.

Nilsson KA, McCann KS. (2016). Interaction strength revisited—clarifying the role of energy flux for food web stability. *Theor Ecol* **9**: 59–71.

Patterson DJ. (2009). Seeing the Big Picture on Microbe Distribution. *Science* **325**: 1506–1507.

Peduzzi P, Concato J, Feinstein AR, Holford TR. (1995). Importance of events per independent variable in proportional hazards regression analysis II. Accuracy and precision of regression estimates. *Journal of Clinical Epidemiology* **48**: 1503–1510.

Shade A, Gregory Caporaso J, Handelsman J, Knight R, Fierer N. (2013). A meta-analysis of changes in bacterial and archaeal communities with time. *ISME J* **7**: 1493–1506.

Smith DP, Peay KG. (2014). Sequence Depth, Not PCR Replication, Improves Ecological Inference from Next Generation DNA Sequencing. *PLOS ONE* **9**: e90234.

Taylor LR. (1961). Aggregation, Variance and the Mean. *Nature* **189**: 732–735.

Thomas MK, Kremer CT, Klausmeier CA, Litchman E. (2012). A Global Pattern of Thermal Adaptation in Marine Phytoplankton. *Science* **338**: 1085–1088.

Tripathi BM, Kim M, Singh D, Lee-Cruz L, Lai-Hoe A, Ainuddin AN, *et al*. (2012). Tropical Soil Bacterial Communities in Malaysia: pH Dominates in the Equatorial Tropics Too. *Microb Ecol* **64**: 474–484.

Trosvik P, de Muinck EJ. (2015). Ecology of bacteria in the human gastrointestinal tract—identification of keystone and foundation taxa. *Microbiome* **3**: 44.

Ulrich W, Gotelli NJ. (2010). Null model analysis of species associations using abundance data. *Ecology* **91**: 3384–3397.

Walter J. and Ley R (2011). The Human Gut Microbiome: Ecology and Recent Evolutionary Changes. *Annual Review of Microbiology* **65**: 411–429.

Weiss S, Van Treuren W, Lozupone C, Faust K, Friedman J, Deng Y, *et al*. (2016). Correlation detection strategies in microbial data sets vary widely in sensitivity and precision. *ISME J* **10**: 1669–1681.

Wootton KL, Stouffer DB. (2015). Many weak interactions and few strong; food-web feasibility depends on the combination of the strength of species' interactions and their correct arrangement. *Theor Ecol* **9**: 185–195.

Worm B, Duffy JE. (2003). Biodiversity, productivity and stability in real food webs. *Trends in Ecology & Evolution* **18**: 628–632.

**Figure 1:** This diagram shows an overview of how our cohesion metrics are calculated, beginning with the relative abundance table and ending with the cohesion values. The relative abundance table shows six samples (S1 indicating "Sample 1", etc.) and a subset of taxa (A, B, C, and Z). First, pairwise correlations are calculated between all taxa, which are entered into the correlation matrix. We then used a null model to account for how the features of microbial datasets might affect correlations, and we subtracted off these values (null model detailed in Fig. 2). For each taxon, we averaged the positive and negative corrected correlations separately and recorded these values as the positive and negative connectedness values. Cohesion values were obtained by multiplying the relative abundance table by the connectedness values. Thus, there are two metrics of cohesion, corresponding to positive and negative values.

**Figure 2:** Microbial data are in the form of relative abundance, and some taxa are much more abundant than others, which are factors that may cause taxa to be spuriously correlated. Thus, we devised a null model to account for the bias that these data features introduce into our metrics. We repeated this process with each taxon as the "focal taxon," which is A in this figure. For each of 200 iterations, we randomized all taxon abundances *besides the focal taxon*. We then calculated correlations between the focal taxon and all other taxa. We recorded the median value of the 200 correlations calculated for each pair of taxa in the median correlation matrix.

**Figure 3:** Comparing the metrics of connectedness obtained from the absolute abundance dataset (x-axes) and the relative abundance dataset (y-axes) shows agreement between the two methods of generating these metrics. Correlations between the metrics are 0.810 (panel A) and 0.741 (panel B). We used separate variables for positive and negative metrics because relativizing the

dataset is expected to differentially affect positive and negative correlations. Solid lines show the fit of linear models.

**Figure 4:** We used our metrics of community cohesion as predictors of the rate of compositional turnover (Bray-Curtis dissimilarity) in the Mendota phytoplankton communities. Negative cohesion was a significant predictor ($p < 1*10^{-20}$) of Bray-Curtis dissimilarity, and the regression explained 46.5% of variation in compositional turnover.

**Figure 5:** We simulated datasets where correlations between taxa were artificially produced by forced correlation to external factors. We calculated cohesion values for the simulated communities to test whether cohesion and Bray-Curtis dissimilarity were strongly related in simulated datasets. The histogram of model adjusted $R^2$ values from our simulations shows that the median adjusted $R^2$ was 0.022 (dashed line), with 95% of values falling below 0.088. For comparison, observed adjusted $R^2$ values ranged from 0.36 to 0.50.

114
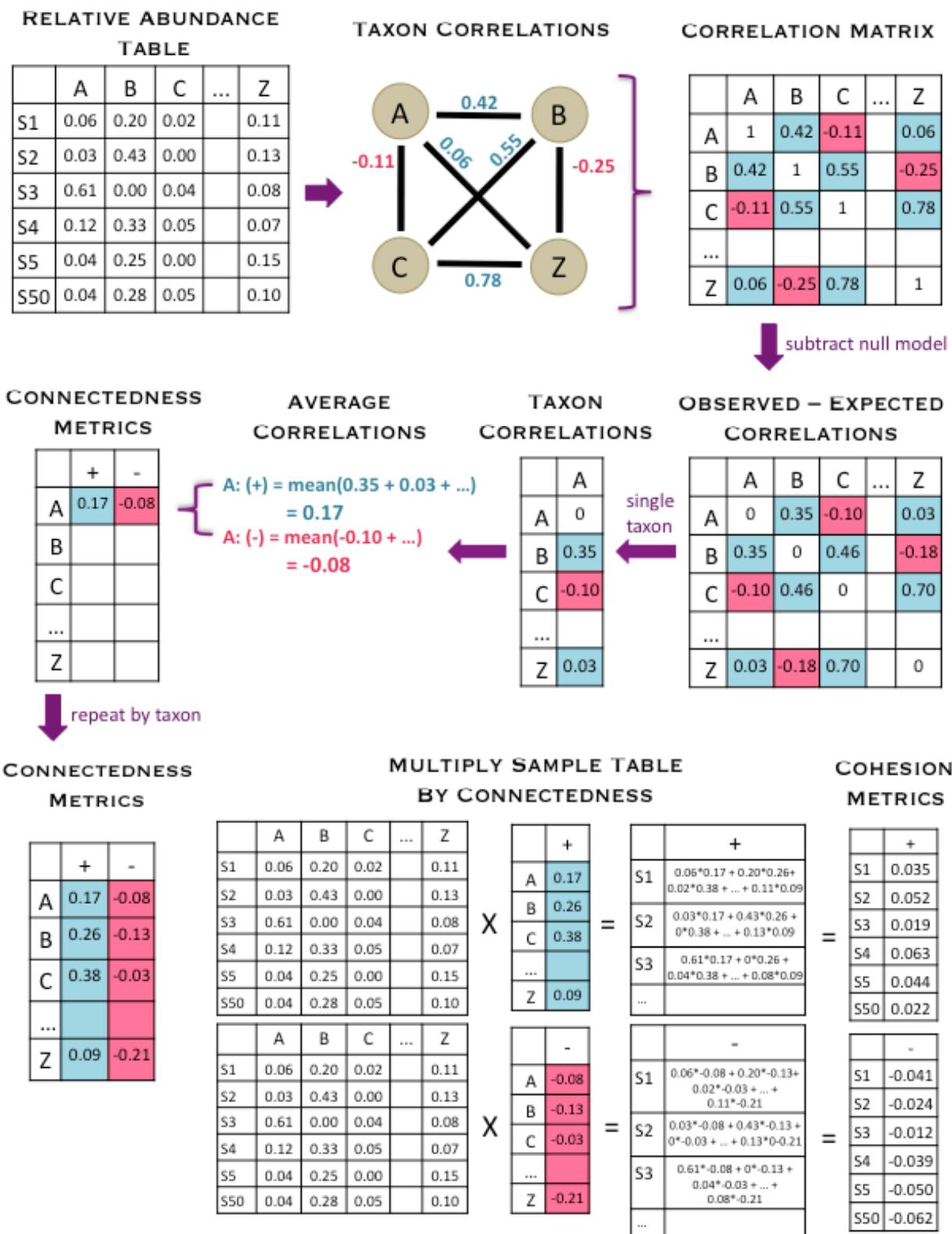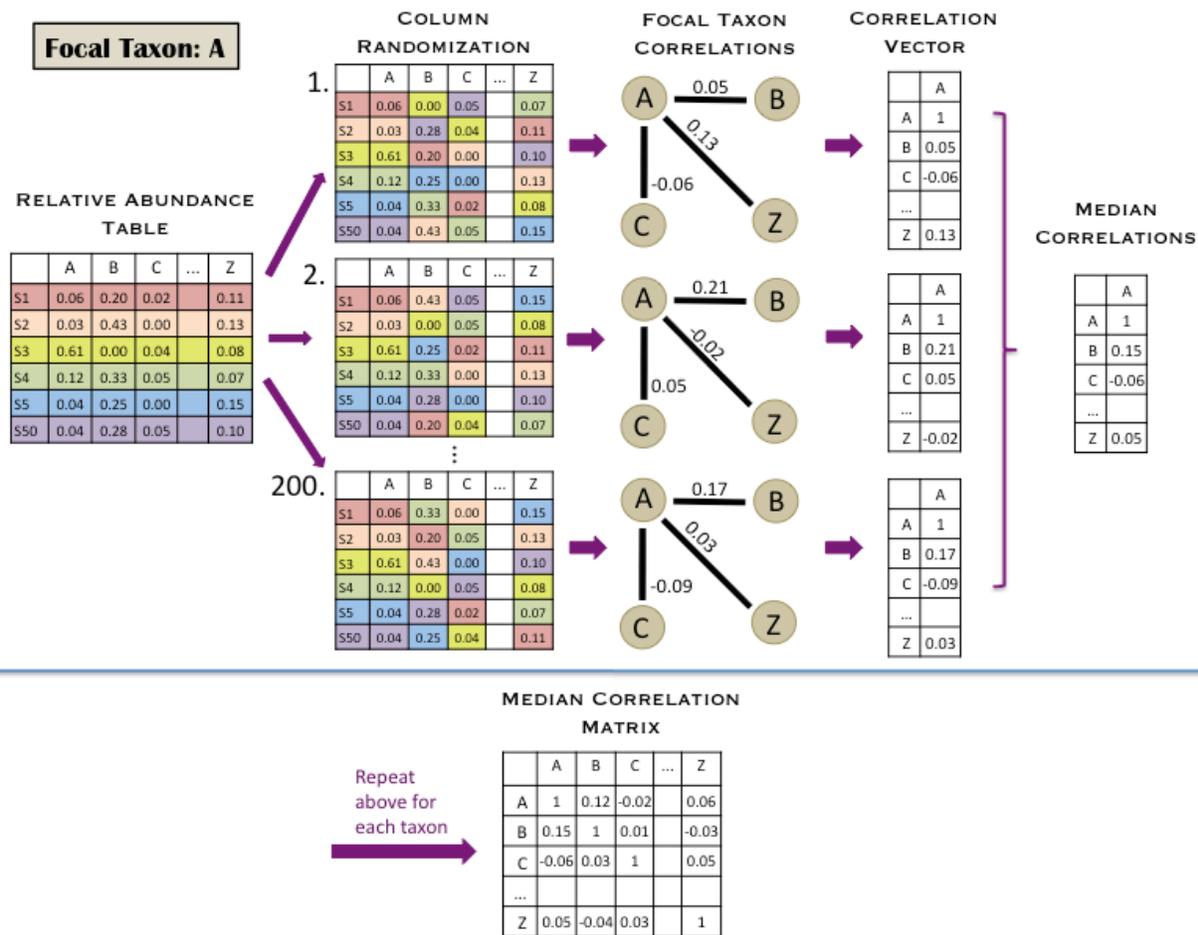
**Figure 1:**
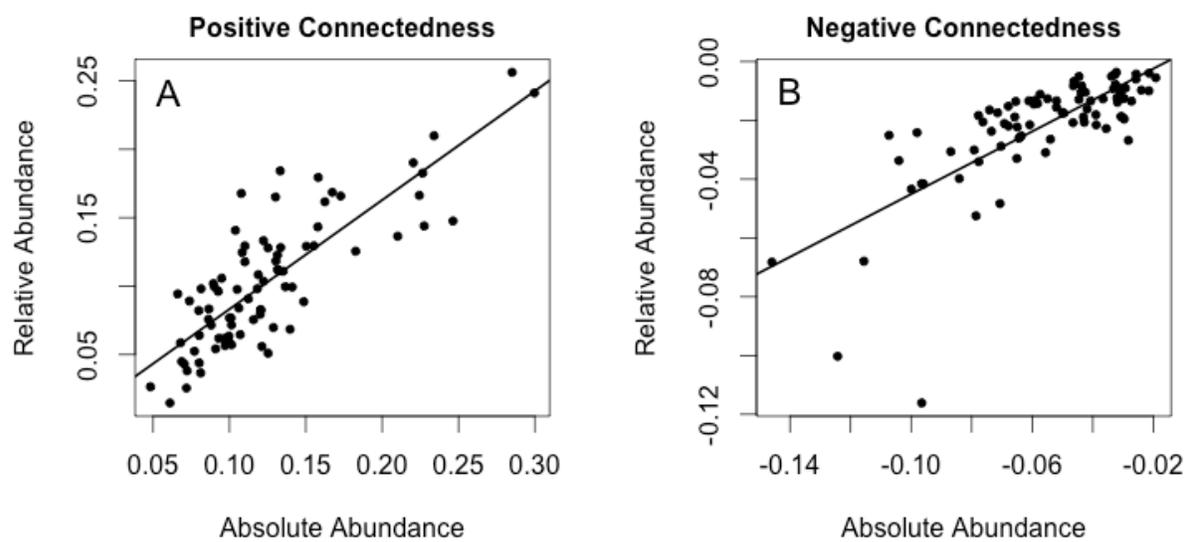
**Figure 2:**

**Figure 3:**

**Figure 4:**
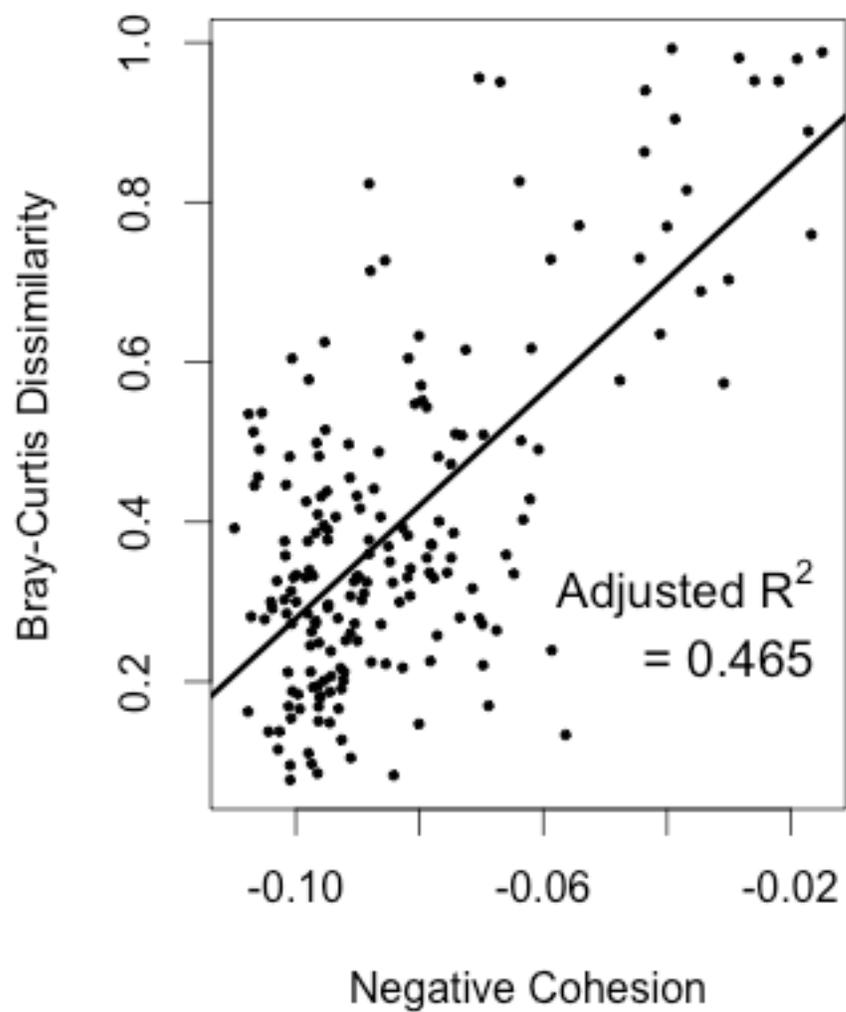
**Figure 5:**
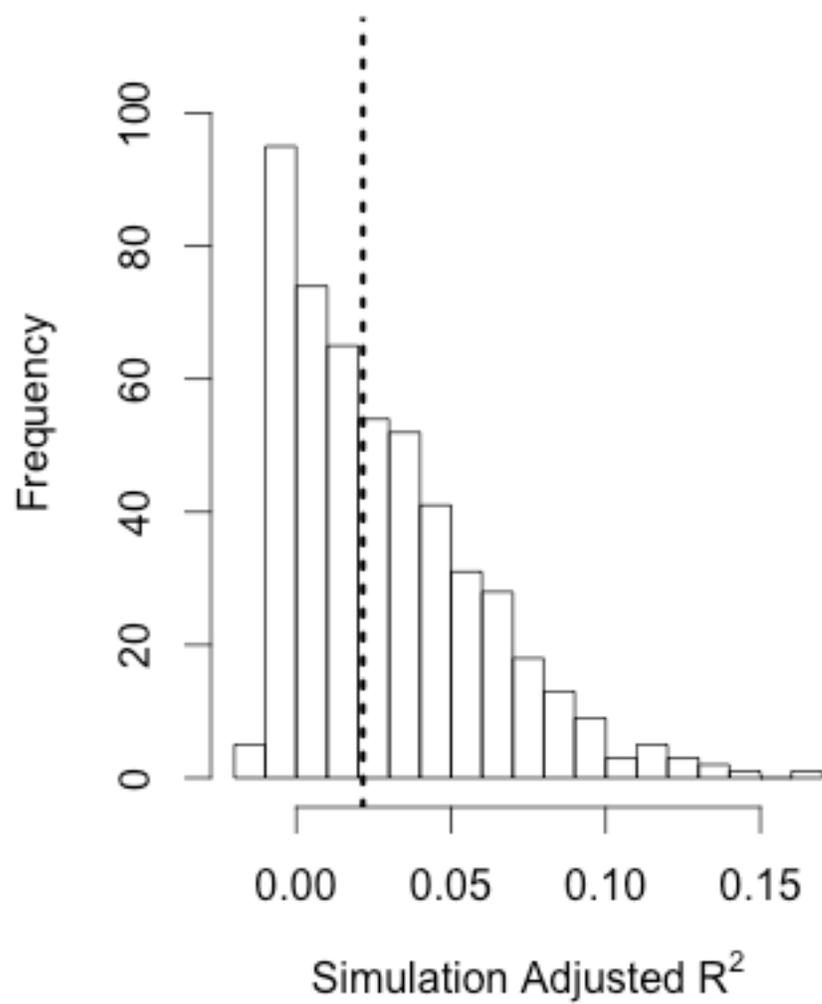
**Table 1: Cohesion predicts community turnover in six long-term time series**

| Lake | Taxon Pers. Cutoff [*] | Model Adjusted $R^2$ | Positive Cohesion P value | Negative Cohesion P value | Positive Cohesion Direction [+] | Negative Cohesion Direction [+] | Days Between Samples | Number of Samples |
|---|---|---|---|---|---|---|---|---|
| Mendota (phyto) | 5% | 0.465 | 0.405 | $< 1*10^{-20}$ | n.s. | Stronger is stabilizing | 36-48 | 186 |
| Monona | 5% | 0.355 | 0.413 | $< 1*10^{-15}$ | n.s. | Stronger is stabilizing | 36-48 | 166 |
| Peter | 10% | 0.357 | 0.062 | $< 1*10^{-3}$ | n.s. | Stronger is stabilizing | 39-45 | 121 |
| Paul | 10% | 0.500 | $< 1*10^{-11}$ | $< 1*10^{-19}$ | Weaker is stabilizing | Stronger is stabilizing | 39-45 | 125 |
| Tuesday | 10% | 0.374 | 0.355 | $< 1*10^{-8}$ | n.s. | Stronger is stabilizing | 39-45 | 72 |
| Mendota (16S) | 5% | 0.378 | 0.0039 | $< 1*10^{-5}$ | Weaker is stabilizing | Stronger is stabilizing | 25-41 | 54 |

\* Stands for "taxon persistence cutoff," which was the minimum proportion of presences across the dataset that we used as a cutoff for including taxa in the connectedness and cohesion metrics. Other cutoffs may give higher model adjusted $R^2$ values (see SOM), but we wanted to use the same cutoff for datasets collected within the same sampling program. We also applied a mean abundance cutoff to the Lake Mendota 16S rRNA gene dataset, where we removed taxa with a mean abundance $< 1*10^{-7}$.

+ These columns indicate the direction of a significant relationship between cohesion and Bray-Curtis dissimilarity. For example, "stronger is stabilizing" means that greater cohesion is related to lower Bray-Curtis dissimilarity. Non-significant relationships are denoted "n.s.".

**Appendix 1: Cohesion R Script**

```
# Online script to generate cohesion metrics for a set of samples
# CMH 26Apr17; cherren@wisc.edu

# User instructions: read in a sample table (in absolute or relative abundance) as object "b".
# If using a custom correlation matrix, read in that matrix at the designated line.
# Run the entire script, and the 4 vectors (2 of connectedness and 2 of cohesion) are generated
for each sample at the end.
# Parameters that can be adjusted include pers.cutoff (persistence cutoff for retaining taxa in
analysis), iter (number of iterations for the null model), tax.shuffle (whether to use taxon shuffle
or row shuffle randomization), and use.custom.cors (whether to use a pre-determined correlation
matrix)

###################create necessary functions####################

#find the number of zeroes in a vector
zero <- function(vec){
  num.zero <- length(which(vec == 0))
  return(num.zero)
}

#create function that averages only negative values in a vector
neg.mean <- function(vector){
  neg.vals <- vector[which(vector < 0)]
  n.mean <- mean(neg.vals)
  if(length(neg.vals) == 0) n.mean <- 0
  return(n.mean)
}

#create function that averages only positive values in a vector
pos.mean <- function(vector){
  pos.vals <- vector[which(vector > 0)]
  p.mean <- mean(pos.vals)
  if(length(pos.vals) == 0) p.mean <- 0
  return(p.mean)
}

##############################################################################
##############################################################################
### Workflow options ####
##############################################################################
##############################################################################

## Choose a persistence cutoff (min. fraction of taxon presence) for retaining taxa in the analysis
pers.cutoff <- 0.10
```

```
## Decide the number of iterations to run for each taxon. (>= 200 is recommended)
# Larger values of iter mean the script takes longer to run
iter <- 200
## Decide whether to use taxon/column shuffle (tax.shuffle = T) or row shuffle algorithm
(tax.shuffle = F)
tax.shuffle <- T
## Option to input your own correlation table
# Note that your correlation table MUST have the same number of taxa as the abundance table.
There should be no empty (all zero) taxon vectors in the abundance table.
# Even if you input your own correlation table, the persistence cutoff will be applied
use.custom.cors <- F


###############################################################################
###############################################################################


# Read in dataset
## Data should be in a matrix where each row is a sample.
b <- read.csv("your_path_here.csv", header = T, row.names = 1)

# Read in custom correlation matrix, if desired. Must set "use.custom.cors" to TRUE
if(use.custom.cors == T) {
  custom.cor.mat <- read.csv("your_path_here.csv", header = T, row.names = 1)
  custom.cor.mat <- as.matrix(custom.cor.mat)
  #Check that correlation matrix and abundance matrix have the same dimension
  print(dim(b)[2] == dim(custom.cor.mat)[2])
}


# Suggested steps to re-format data. At the end of these steps, the data should be in a matrix "c"
where there are no empty samples or blank taxon columns.
c <- as.matrix(b)
c <- c[rowSums(c) > 0, colSums(c) > 0]

# Optionally re-order dataset to be in chronological order. Change date format for your data.
#c <- c[order(as.Date(rownames(c), format = "%m/%d/%Y")), ]

# Save total number of individuals in each sample in the original matrix. This will be 1 if data are
in relative abundance, but not if matrix c is count data
rowsums.orig <- rowSums(c)

# Based on persistence cutoff, define a cutoff for the number of zeroes allowed in a taxon's
distribution
zero.cutoff <- ceiling(pers.cutoff * dim(c)[1])

# Remove taxa that are below the persistence cutoff
d <- c[ , apply(c, 2, zero) < (dim(c)[1]-zero.cutoff) ]
```

```
# Remove any samples that no longer have any individuals, due to removing taxa
d <- d[rowSums(d) > 0, ]

#If using custom correlation matrix, need to remove rows/columns corresponding to the taxa
below persistence cutoff
if(use.custom.cors == T){
  custom.cor.mat.sub <- custom.cor.mat[apply(c, 2, zero) < (dim(c)[1]-zero.cutoff), apply(c, 2,
zero) < (dim(c)[1]-zero.cutoff)]
}

# Create relative abundance matrix.
rel.d <- d / rowsums.orig
# Optionally, check to see what proportion of the community is retained after cutting out taxa
hist(rowSums(rel.d))

# Create observed correlation matrix
cor.mat.true <- cor(rel.d)

# Create vector to hold median otu-otu correlations for initial otu
med.tax.cors <- vector()

# Run this loop for the null model to get expected pairwise correlations
# Bypass null model if the option to input custom correlation matrix is TRUE
if(use.custom.cors == F) {
ifelse(tax.shuffle, {
  for(which.taxon in 1:dim(rel.d)[2]){

    #create vector to hold correlations from every permutation for each single otu
    ## perm.cor.vec.mat stands for permuted correlations vector matrix
    perm.cor.vec.mat <- vector()

    for(i in 1:iter){
      #Create empty matrix of same dimension as rel.d
      perm.rel.d <- matrix(numeric(0), dim(rel.d)[1], dim(rel.d)[2])
      rownames(perm.rel.d) <- rownames(rel.d)
      colnames(perm.rel.d) <- colnames(rel.d)

      #For each otu
      for(j in 1:dim(rel.d)[2]){
        # Replace the original taxon vector with a permuted taxon vector
        perm.rel.d[, j ] <- sample(rel.d[ ,j ])
      }

      # Do not randomize focal column
      perm.rel.d[, which.taxon] <- rel.d[ , which.taxon]
```

```
    # Calculate correlation matrix of permuted matrix
    cor.mat.null <- cor(perm.rel.d)

    # For each iteration, save the vector of null matrix correlations between focal taxon and other
taxa
    perm.cor.vec.mat <- cbind(perm.cor.vec.mat, cor.mat.null[, which.taxon])

  }
  # Save the median correlations between the focal taxon and all other taxa
  med.tax.cors <- cbind(med.tax.cors, apply(perm.cor.vec.mat, 1, median))

  # For large datasets, this can be helpful to know how long this loop will run
  if(which.taxon %% 20 == 0){print(which.taxon)}
 }
} , {
  for(which.taxon in 1:dim(rel.d)[2]){

  #create vector to hold correlations from every permutation for each single otu
  ## perm.cor.vec.mat stands for permuted correlations vector matrix
  perm.cor.vec.mat <- vector()

  for(i in 1:iter){
    #Create duplicate matrix to shuffle abundances
    perm.rel.d <- rel.d

    #For each taxon
    for(j in 1:dim(rel.d)[1]){
      which.replace <- which(rel.d[j, ] > 0 )
      # if the focal taxon is greater than zero, take it out of the replacement vector, so the focal
abundance stays the same
      which.replace.nonfocal <- which.replace[!(which.replace %in% which.taxon)]

      #Replace the original taxon vector with a vector where the values greater than 0 have been
randomly permuted
      perm.rel.d[j, which.replace.nonfocal] <- sample(rel.d[ j, which.replace.nonfocal])
    }

    # Calculate correlation matrix of permuted matrix
    cor.mat.null <- cor(perm.rel.d)

    # For each iteration, save the vector of null matrix correlations between focal taxon and other
taxa
    perm.cor.vec.mat <- cbind(perm.cor.vec.mat, cor.mat.null[, which.taxon])

  }
  # Save the median correlations between the focal taxon and all other taxa
```

```
    med.tax.cors <- cbind(med.tax.cors, apply(perm.cor.vec.mat, 1, median))

    # For large datasets, this can be helpful to know how long this loop will run
    if(which.taxon %% 20 == 0){print(which.taxon)}
  }
 }
)
}
```

```
# Save observed minus expected correlations. Use custom correlations if use.custom.cors =
TRUE
ifelse(use.custom.cors == T, {
  obs.exp.cors.mat <- custom.cor.mat.sub}, {
    obs.exp.cors.mat <- cor.mat.true - med.tax.cors
  }
)
```

```
diag(obs.exp.cors.mat) <- 0
```

```
####
#### Produce desired vectors of connectedness and cohesion
```

```
# Calculate connectedness by averaging positive and negative observed - expected correlations
connectedness.pos <- apply(obs.exp.cors.mat, 2, pos.mean)
connectedness.neg <- apply(obs.exp.cors.mat, 2, neg.mean)
```

```
# Calculate cohesion by multiplying the relative abundance dataset by associated connectedness
cohesion.pos <- rel.d %*% connectedness.pos
cohesion.neg <- rel.d %*% connectedness.neg
```

```
####
#### Combine vectors into one list and print
output <- list(connectedness.neg, connectedness.pos, cohesion.neg, cohesion.pos)
names(output) <- c("Negative Connectedness", "Positive Connectedness", "Negative Cohesion",
"Positive Cohesion")
```

```
print(output)
```

**Appendix 2: readme for Cohesion R Script**

script maintainer: Cristina Herren, cristina.herren@gmail.com

*While developing the cohesion analysis, we tested our workflow on many different datasets.*

*These are some suggested best practices and diagnostics for using the cohesion metrics.*

Suggestions for setting abundance and persistence thresholds:

- The cohesion R script includes a parameter (pers.cutoff) to exclude taxa from the analysis if they are present in fewer than a specified proportion of samples. The purpose of this parameter is to exclude taxa for which reliable correlations cannot be calculated. As a guideline, taxa should be present in at least 5-10 samples to be included in the analysis. Thus, if your dataset has 50 samples, the pers.cutoff parameter might initially be set to 0.1, as to exclude taxa present in fewer than 5 samples.

- For samples with deep sequencing, rare taxa (taxa with very low abundances) are present in a large proportion of samples. In this case, it may be optimal to cut out taxa from the analysis based on mean abundance. As a guideline, the abundance threshold might be initially set to exclude, on average, 5-10% of the community. In our analyses, this approach dramatically reduced the number of taxa included (thereby also making the script run faster) while retaining the vast majority of the community.

- When using cohesion as a predictor variable in analyses, we found that analysis results were often qualitatively similar across a range of abundance and persistence thresholds (see sensitivity analysis SOM). Thus, we suggest trying various thresholds and selecting final parameter values within the range of values where results are stable.

Suggestions for correcting or importing the correlation matrix:

- While creating the cohesion workflow, we tested dozens of null models for correcting correlations between taxa. We selected a final version based on what worked well for a variety of different datasets. However, we imagine that the default null models included in the script might not be ideal for every dataset, because microbial datasets vary in richness and evenness. You can test the influence of the null model on the cohesion metrics by importing the true (uncorrected) correlation matrix as a custom correlation matrix. This will bypass the null model and calculate connectedness on the uncorrected correlation matrix.

- We found that the row shuffle null model worked better as sample evenness increased. Conversely, the column shuffle null model worked better when there were taxa that consistently comprised > 20% of the community. In these cases, it was important to maintain taxon mean abundances in the null model, which only occurs in the column shuffle null model.

- If importing a custom correlation matrix, we urge users to consider the aim of the analysis used to generate the custom matrix. Correlation methods that are intended to determine significance may not yield appropriate correlation matrices. For example, in Local Similarity Analysis (LSA), LS scores do not equate to significance; a lower LS score can be more significant than a higher LS score. This is not a deficit of the Local Similarity method, because the aim of LSA is to identify significant pairwise correlations. However, the differing objectives of these two analyses (LSA and cohesion) mean that the matrix produced from LSA may not be an appropriate custom matrix.

Types of datasets appropriate for the cohesion metrics

- Time series datasets from a single location

- Time series datasets from multiple similar locations (e.g. water samples from across a watershed or from many groundwater wells)

- Spatial datasets (e.g. soil samples across a landscape)

- Host-associated microbiome samples from a single individual/organism (e.g. stool samples from a single human subject)

- Host-associated microbiome samples from multiple individuals/organisms within the same population (e.g. stool samples from many human subjects)

- Experimental datasets (e.g. calculating cohesion metrics or the relationship between cohesion and a response variable in two or more experimental treatments)

Types of datasets that may be problematic for the cohesion metrics

- Time series collected at high frequencies (such as a daily time scale)

  o An implicit assumption of this analysis is that there is sufficient time between samples for populations to change enough in relative abundance such that the real population change can be distinguished from background/methodological noise in the data. This assumption may not be met at a daily sampling interval.

  o Another problem that may arise in time series at high frequency is autocorrelation. Autocorrelation may inflate correlations between taxa. One possible way to account for autocorrelation would be to calculate connectedness using the first differences of the dataset. Correlations on the first differences would indicate whether changes in two populations are synchronous.

- Datasets containing samples from different sites where most OTUs are not shared.

  o OTUs being present at one site but absent at others may generate spurious correlations between taxa.

Diagnostic tests when cohesion is used as a predictor in a regression

- Test for autocorrelation in residuals. As a simple diagnostic, plot sequential residuals against one another. Test for significant correlation between sequential residuals.

- If analyzing a time series, test for trends in residuals over time.

- If analyzing a time series, split the dataset into two parts corresponding the first half of the time series and the second half of the time series. Fit separate regressions for the two halves, and test whether the slope estimates for cohesion are significantly different.

- Cross validation: randomly split dataset into two halves many times. Use one half to train the model, and the other half to calculate the error in predicted values from the linear model. Compare the residual error in the fitted model to the error in the predicted values.

- Test the significance and model $R^2$ value of cohesion metrics when cohesion metrics are used to predict dynamics of random samples. To do this, randomize the order of your samples. Conduct the same analysis (such as a regression of cohesion predicting Bray-Curtis dissimilarity) using randomly ordered samples. Repeat many times to get a distribution of results. Cohesion should be more significant when used to predict dynamics in ordered samples, versus randomized samples.

- The most abundant taxon in microbial communities can comprise more than 10% of the entire community. To test how strongly the most abundant taxon contributes to cohesion values, manually set the connectedness values of the most abundant taxon to zero. Re-calculate cohesion, and re-run the analysis using these new cohesion metrics.

## We suggest not rarefying datasets before calculating cohesion

- We found that rarefying decreases the strength of pairwise correlations. We tested the effects of rarefying on pairwise correlations using a dataset of bacterial samples from bog lakes in northern Wisconsin (available at the Earth Microbiome Project, study ID 1288). This dataset contained replicate samples at many time points. We calculated correlations between the same OTU in replicated samples in the non-rarefied dataset and in a rarefied dataset. These correlations should be close to 1. We found that rarefying consistently decreased the magnitude of these strong correlations. We also found that rarefying increased the standard deviation of OTUs due to the stochasticity of rarefying. In the equation for Pearson correlations, the denominator is the product of the standard deviations of the two populations. Thus, the mathematical reason for the decreased correlations is the increase in the standard deviations of the OTUs without a compensating increase in the covariance of the two populations.

## Miscellaneous

- The null models implemented in the cohesion R script are stochastic due to random sampling. To generate reproducible results, manually set a seed from a predetermined vector before each randomization.

- When running the script on a datasets where correlations for 500 taxa were calculated, the script took 30-60 minutes to run. The run time is directly proportional to the number of iterations (iter). We found that using iter = 40 was sufficient during parameter optimization (i.e. while determining the persistence cutoff and type of null model to use), but that a larger value (iter = 200) was best to generate final results.

- If using DNA sequencing data, try clustering the sequences based on different percent similarity cutoffs.

**Appendix 3: R Script Used for Simulations of Correlated Taxa**

```
# Make script to artificially correlate taxa to one another to investigate whether results of
workflow and analysis are an artifact of correlated datasets
# CMH 30Nov16

library(reshape)
library(Hmisc)
library(plyr)
library(vegan)
library(AICcmodavg)
library(car)

#########create necessary functions ##########

zero <- function(vec){
  num.zero <- length(which(vec == 0))
  return(num.zero)
}

#create function that averages only negative values in a vector
neg.mean <- function(vector){
  neg.vals <- vector[which(vector < 0)]
  n.mean <- sum(neg.vals) / length(neg.vals)
  ifelse(length(neg.vals) == 0, return(0), return(n.mean) )
}

#create function that averages only positive values in a vector
pos.mean <- function(vector){
  pos.vals <- vector[which(vector > 0 & vector < 1)]
  p.mean <- sum(pos.vals) / length(pos.vals)
  ifelse(length(pos.vals) == 0, return(0), return(p.mean) )
}

#Import code that creates a vector with desired correlation to other vector
## Found on Stack Exchange by user caracal
gen.cor.vec <- function(vec.orig, cor){
  n    <- length(vec.orig)    # length of vector
  rho  <- cor                 # desired correlation = cos(angle)
  theta <- acos(rho)          # corresponding angle
  x1   <- vec.orig       # fixed given data
  x2   <- rnorm(n, 2, 1)      # new random data
  X    <- cbind(x1, x2)        # matrix
  Xctr <- scale(X, center=TRUE, scale=FALSE)   # centered columns (mean 0)

  Id   <- diag(n)                      # identity matrix
```

```
Q    <- qr.Q(qr(Xctr[ , 1, drop=FALSE]))     # QR-decomposition, just matrix Q
P    <- tcrossprod(Q)         # = Q Q'      # projection onto space defined by x1
x2o  <- (Id-P) %*% Xctr[ , 2]               # x2ctr made orthogonal to x1ctr
Xc2  <- cbind(Xctr[ , 1], x2o)              # bind to matrix
Y    <- Xc2 %*% diag(1/sqrt(colSums(Xc2^2)))  # scale columns to length 1

x <- Y[ , 2] + (1 / tan(theta)) * Y[ , 1]     # final new vector
cor(x1, x)
return(x)
}


###############################################################
### Parameter Inputs ####
###############################################################

#how many samples in matrix
num.samples <- 187
#choose uniform or t distribution for correlations
uniform <- F
#choose a cutoff for presence for simulating a taxon's abundance in the generated dataset.
## It's computationally intensive to generate many rare taxa that get cut in the pipeline anyway
zero.prop <- .015
## Choose a persistence cutoff for taxa in the workflow with simulated datsets
zero.prop.sim <- .05
#set a proportion of the simulated dataset to 0, as detection limit
det.limit <- .5

#Read in dataset to use as template as p
## For example:
#p <- read.csv("~/Desktop/MendotaPhytoTable.csv", header = T, row.names = 1)

c <- as.matrix(p)
c <- c[rowSums(c) > 0, colSums(c) > 0]
c <- c[order(as.Date(rownames(c), format = "%m/%d/%Y")), ]

#define zero cutoff based on rows of c
zero.cutoff <- ceiling(zero.prop * dim(c)[1])

c <- c[ , apply(c, 2, zero) < (dim(c)[1]-zero.cutoff) ]
c <- c[rowSums(c) > 0, ]

#Save curated count and relative datasets
orig.count <- c
orig.rel <- sweep(orig.count, 1, rowSums(orig.count), "/")
```

```
################################################################################
## Run a loop to simulate run analysis on 1000 datasets
################################################################################
loop.iter <- 500

#Create a matrix to hold coefficient estimates for regression modeling Bray-Curtis Dissimilarity
sim.loop.results <- matrix(numeric(0), 5, loop.iter)
rownames(sim.loop.results) <- c("pos.cohesion est", "pos.cohesion t val", "neg.cohesion est",
"neg.cohesion t val", "model.R2")

for(loop in 1:loop.iter){

#Specify number of time steps for environmental vectors by number of samples
time.vec <- seq(1, num.samples, 1)

#Define 4 environmental driver vectors
envs1 <- arima.sim(list(order = c(1,1,0), ar = 0.9), n = max(time.vec-1), sd = 1)
envs2 <- arima.sim(list(order = c(1,1,0), ar = 0.9), n = max(time.vec-1), sd = 1)
envs3 <- arima.sim(list(order = c(1,1,0), ar = 0.9), n = max(time.vec -1), sd = 1)
envs4 <- arima.sim(list(order = c(1,0,0), ar = 0.1), n = max(time.vec), sd = 1)

envs.mat <- rbind(envs1, envs2, envs3, envs4)
cor(t(envs.mat))

#Define means of taxa. Use means of original dataset
otu.means <- 1 * apply(orig.count, 2, mean)

#Create a loop to generate a matrix of OTU abundances
#First, choose what environmental drivers each OTU will be correlated with
num.otus <- length(otu.means)
otu.env.drivers <- sample(c(1, 2, 3, 4), num.otus, replace = T)

#Create a vector to determine how strongly taxa will be correlated with drivers
t.cors <- rt(num.otus, 4, 0) #use a t distribution because correlations are fat-tailed
t.cors <- t.cors / max(abs(t.cors) + .1) #bound correlations by -1 and 1
ifelse(uniform, otu.env.cors <- runif(num.otus, -.8, .8), otu.env.cors <- t.cors)

#Create vectors to hold OTU abundances
gen.cors <- vector()
otu.abun <- vector()

for(i in 1:num.otus){
  #generate a vector that is correlated with the randomly designated environmental driver
  cor.abun <- gen.cor.vec(envs.mat[otu.env.drivers[i], ], otu.env.cors[i])

  #make generated values positive
```

```
cor.abun.pos <- cor.abun + abs(min(cor.abun))
#give the abundance vector a standard deviation of 1
cor.abun.pos.sd1 <- cor.abun.pos / ( sd(cor.abun.pos) )

#multiply by mean value
scaled.abun.pos <- cor.abun.pos.sd1 * otu.means[i]

#draw individual abundances from lognormal distribution
## add 1 to mean parameter before taking log, so it cannot be negative.
## introuce between-taxon variability by drawing sd parameter from a normal distribution
log.abun <- rlnorm(length(scaled.abun.pos), log(scaled.abun.pos + 1), sd = rnorm(1, 1, .2))

#Save vectors generated from correlation function
gen.cors <- cbind(gen.cors, cor.abun)
sd1.abun <- cor.abun / sd(cor.abun)
otu.abun <- cbind(otu.abun, log.abun)
}

## Find 50% quantile for current matrix and set to zero, as detection limit
quant.zero <- sort(otu.abun)[length(sort(otu.abun)) * det.limit]

## Replace all abundance values below "detection limit" with 0
otu.abun.noise <- otu.abun
otu.abun.noise[otu.abun < quant.zero] <- 0
otu.abun.noise <- otu.abun.noise[rowSums(otu.abun.noise) > 0, colSums(otu.abun.noise) > 0]

#Relativize datasets
otu.rel <- sweep(otu.abun.noise, 1, rowSums(otu.abun.noise), "/")
colnames(otu.rel) <- seq(1, dim(otu.rel)[2], 1)

###############################################################################
## Start analysis pipeline
###############################################################################

ab.raw <- otu.rel[rowSums(otu.rel) > 0, colSums(otu.rel) > 0]
ab.raw <- as.matrix(ab.raw)

##############################################################
#create new matrix so original input is saved
d.sub <- ab.raw

#take out OTUs with fewer occurrences than the presence cutoff
zero.cutoff.sim <- ceiling(zero.prop.sim*dim(d.sub)[1])
d.sub <- d.sub[, apply(d.sub, 2, zero) < (dim(d.sub)[1]-zero.cutoff.sim) ]
d.sub <- as.matrix(d.sub)
d.sub <- d.sub[rowSums(d.sub) > 0, colSums(d.sub) > 0]
```

```
cor.mat.dsub <- cor(d.sub)
hist(cor.mat.dsub)

#define number of randomizations from which to get expected cors
perm.cor <- 100

med.otu.cors <- vector()

 for(m in 1:dim(d.sub)[2]){
  which.otu <- m

  #create vector to hold correlations from every permutation for each single otu
  perm.cor.vec.mat <- vector()

  for(i in 1:perm.cor){
   d.sub2 <- d.sub
   for(j in 1:dim(d.sub)[2]){
     d.sub2[, j ] <- sample(d.sub[ ,j ])
    }
   #replace focal column with original column
   d.sub2[, which.otu] <- d.sub[ , which.otu]

   cor.mat.dsub.null <- cor(d.sub2)

   perm.cor.vec.mat <- cbind(perm.cor.vec.mat, cor.mat.dsub.null[,m])

  }
  med.otu.cors <- cbind(med.otu.cors, apply(perm.cor.vec.mat, 1, median))

  #if(m %% 20 == 0){print(m)}
 }


#the vector med.otu.cors holds the median null model otu-otu correlations
diag(med.otu.cors) <- 0
#get observed - expected individual correlations
obs.exp.cors.each <- cor.mat.dsub - med.otu.cors
diag(obs.exp.cors.each) <- 0

#Calculate positive and negative connectedness
connectedness.pos <- apply(obs.exp.cors.each, 2, pos.mean)
connectedness.neg <- apply(obs.exp.cors.each, 2, neg.mean)



###############################################################################
```

```
################################################################################

#look at distribution of BC dissimilarities of samples
bc.diss <- as.matrix(vegdist(ab.raw, method = "bray"))

#test turnover in WHOLE community, not just those included in the subset
#pull out turnover from time t to t+1
  bc.vec <- vector()
  for(i in 1:dim(ab.raw)[1]-1){
    bc.vec[i] <- bc.diss[i, i+1]
  }

cohesion.pos <- as.matrix(d.sub) %*% connectedness.pos
cohesion.neg <- as.matrix(d.sub) %*% connectedness.neg

cohesion.summary <- summary(lm(bc.vec ~ cohesion.pos[-length(cohesion.pos)] +
cohesion.neg[-length(cohesion.neg)]))

#Save results of multiple regression
sim.loop.results[1, loop] <- coef(cohesion.summary)[2, 1]
sim.loop.results[2, loop] <- coef(cohesion.summary)[2, 3]
sim.loop.results[3, loop] <- coef(cohesion.summary)[3, 1]
sim.loop.results[4, loop] <- coef(cohesion.summary)[3, 3]
sim.loop.results[5, loop] <- cohesion.summary$adj.r.squared

#This can be useful to estimate how long the program will run
if(loop %% 50 == 0)(print(loop))

}

sim.loop.results.mat <- sim.loop.results
sim.loop.results <- sim.loop.results.mat[, !is.na(colSums(sim.loop.results.mat))]

#Look at distribution of significance values (t values) for pos and neg cohesion
par(mfrow = c(2,1))
hist(sim.loop.results[2,], main = rownames(sim.loop.results)[2])
mean(sim.loop.results[2,] > 3.35 | sim.loop.results[2,] < -3.35 )
hist(sim.loop.results[4,], main = rownames(sim.loop.results)[4])
mean(sim.loop.results[4,] > 3.35 | sim.loop.results[4,] < -3.35 )

#plot r2 observed and theoretical R2
par(mfrow = c(1, 1))
hist(sim.loop.results[5,], breaks = 20, xlab = expression("Adjusted R"^2), main =
expression("Compositional Turnover Model R" ^2 ))
abline(v = median(sim.loop.results[5,]) , lwd = 3, lty = 2)
```

```
#Report median simulation R2 value and 95% boundary
median(sim.loop.results[5,])
sort(sim.loop.results[5,])[.95*dim(sim.loop.results)[2]]


#write.csv(sim.loop.results, "~/Desktop/MethodsSimTable.csv")
```

**Appendix 4: Sensitivity Analysis and Alternate Null Models**

<u>Sensitivity Analysis Metadata</u>

Interpretation of column headers in sensitivity analysis spreadsheet:

LakeID ➔ name of lake analyzed

PersCutoff ➔ taxa below this persistence cutoff (defined as the proportion of samples present across the entire dataset) were not included in calculating connectedness or cohesion metrics

ColShuffle ➔this variable is TRUE or FALSE, based on which null model was used. If TRUE, the null model described in the main text was used (i.e. abundance values were shuffled by taxon, which appear as columns in the dataset). If FALSE, abundance values were shuffled by sample (appearing as rows in the datset); abundance values of all *present* taxa were randomized.

DaysElapseMin ➔ minimum number of days between paired samples for which Bray-Curtis dissimilarities were calculated

DaysElapseMax ➔ maximum number of days between paired samples for which Bray-Curtis dissimilarities were calculated

ModelR2 ➔ model $R^2$ value of the multiple regression predicting Bray-Curtis dissimilarity as a function of positive cohesion and negative cohesion

PosCohEst ➔ model coefficient estimated for positive cohesion in the multiple regression

PosCohPval ➔ p value for positive cohesion from the multiple regression

NegCohEst ➔ model coefficient estimated for negative cohesion in the multiple regression

NegCohPval ➔ p value for negative cohesion from the multiple regression

NumSamples ➔ number of pairs of samples (and associated Bray-Curtis dissimilarity values) included in the analysis, based on the minimum and maximum numbers of days elapsed

<u>Description of alternate null models investigated</u>

As described in the main text, we found that taxon connectedness values were related to taxon average abundance and persistence across the dataset. This was particularly true for negative correlations, with more abundant taxa having stronger negative correlations. When developing null models, we tried approaches that maintained either taxon distributions (column sums) or sample distributions (row sums).

The benefit of models that maintain column sums (such as the null model presented in the text) is that no taxon would be assigned an abundance value in the null model that was not possible for them to attain. Furthermore, it would keep mean taxon abundances the same, which we identified as one possible contributor to average taxon correlations. The benefit of models that maintain row sums is that negative dependencies between taxa still exist in this null model. Thus, if one taxon "bloomed" in a certain sample, and all other taxa decreased in relative abundance, this effect would be captured in the null model.

The alternative model implemented in the cohesion R script maintains row sums and taxon persistence (fraction of presences across the dataset) during the randomization step. The sensitivity analysis shows results of the analyses presented in the main text when the row-shuffle model is used instead of the taxon/column shuffle null model. In the row-shuffle null model, abundances of taxa that are present in a given sample are randomly assigned. In some instances, this null model performed better than the model used in the main text; this occurred when the model $R^2$ of the cohesion variables in the multiple regression was higher when using the row-shuffling null model versus the column-shuffling null model. However, these instances were uncommon across the five datasets, and the improvement in predictive power when using the row-shuffling null model was often not statistically significant (as determined by comparing model AICc values). Additionally, there were many times where the row-shuffling null model

performed much worse than the column-shuffling null model, particularly in Peter, Paul, and Tuesday lakes. However, the row-shuffling null model gave a higher model $R^2$ in the Bray-Curtis dissimilarity analysis than the column-shuffling null model (see Lake Mendota 16S analysis supplementary material).

We tried many variations on these two broad null model types (row shuffling and column/taxon shuffling) but rejected many options based on poor preliminary results. Here we give a list of alternate algorithms tested while developing this workflow. Many of these options were tried in combination, leading to several dozen null models tested.

- Instead of recording median correlations between each pair of taxa in the null model, we recorded median connectedness values for each taxon. We then subtracted the null model estimates of connectedness from the observed values of connectedness to arrive at the "corrected" connectedness values.

- We tried using the mean of the null model correlations as the "expected" correlation, instead of the median. This option gave very similar results to the final models used.

- We tried an alternate method of calculating connectedness metrics from the correlation matrix; instead of using the average positive or negative correlations, we divided the sum of all positive or negative correlations by the total number of pairwise correlations.

- We tried a variation on the column-shuffling null model, where only abundance values greater than zero were randomized (i.e. only values of present taxa were sampled).

- We tried a variation on the row-shuffling null model where all abundance values were randomized (not just present taxa).

- We tried using raw (uncorrected) correlations when calculating connectedness values. In many instances, there was no strong difference in outcome when using either uncorrected

or corrected correlations. However, we expect that using uncorrected correlations would be problematic in more uneven datasets.

- We tried randomizing abundance values of the focal taxon, rather than all other taxa. With enough iterations, there was no difference between this null model and the one described in the main text.

- We tried many alternative algorithms for calculating correlations between taxa. These included Spearman correlations, calculating the log variance of the proportions of the two taxa, and transforming the abundances (log transform or square root transform) before calculating Pearson correlations.

- Our connectedness metrics average all positive or negative correlations, regardless of strength. We instead tried implementing a correlation cutoff value, such that only correlations stronger than the cutoff would be included in the connectedness metrics. However, we found that this method gave lower explanatory power in the regression of Bray-Curtis dissimilarity vs. cohesion.

**Appendix 5: Environmental Model of Lake Mendota Phytoplankton**

The purpose of this model was to assess how well the available environmental data can model the rate of compositional turnover in Lake Mendota phytoplankton. Although this analysis almost certainly excludes important variables (either unmeasured, or interactions not included), it gives an estimate of the predictive power of a moderately well-informed model using environmental data. Furthermore, the North Temperate Lakes LTER is rare in its scope and length of data collection; the quality and regularity of sampling of these data likely enables greater explanatory power than in other phytoplankton datasets.

We downloaded environmental data from the North Temperate Lakes LTER website (https://lter.limnology.wisc.edu/datacatalog/search). The available datasets for Lake Mendota that span the years 1995 – 2013 are: major ion concentrations, nutrient concentrations (nitrogen, phosphorus, silica), carbon concentrations, pH, alkalinity, chlorophyll and phaeophytin concentrations, Secchi depth, air temperature, water temperature, dissolved oxygen concentration, dissolved oxygen saturation, and instantaneous wind speed, wave height, and cloud cover.

Several variables were excluded from our analysis because they were not appropriate for this model. We chose not to analyze instantaneous wind speed, wave height, or cloud cover, reasoning that these variables would not influence turnover in phytoplankton communities on the scale of 5-7 weeks. There were a large number of missing values in the pH, alkalinity, and carbon datasets, which is why these variables were excluded. Major ion concentrations were measured too infrequently to be included in the model. Chlorophyll and phaeophytin data were missing from early 2002 through the end of 2004, and so were excluded as variables. The following variables were measured with sufficient regularity as to be used as predictors in the

model: water temperature, air temperature, $NO_3 + NO_2$ concentrations, $NH_4$ concentrations, total nitrogen concentrations, dissolved reactive phosphorus concentrations, total phosphorus concentrations, dissolved silica concentrations, Secchi depth, dissolved oxygen concentration, dissolved oxygen saturation. These variables were all measured on the same dates as phytoplankton community samples were taken.

The metadata of the Lake Mendota phytoplankton samples indicates that water samples were pooled from the top 0-8m of the water column. Thus, for variables where depth was indicated (e.g. water temperature, dissolved oxygen concentrations), we removed values from depths below 8m. Then, if multiple values were still present, we averaged the values across depths 0-8m. These averages were used as the predictor for that sample date.

The dependent variable in this model was the Bray-Curtis dissimilarity between each of the 186 pairs of communities where the time elapsed between sample dates was 36-48 days. These are the same 186 paired communities used in the regression where cohesion was the predictor. We matched each Bray-Curtis dissimilarity value with the environmental variables taken on the same date as the first of the paired communities. Thus, each environmental variable lad a length of 186, although values were missing for some variables. We created vectors for the change in the environmental drivers by subtracting the value of the environmental driver on the first sample date from the value of the environmental driver on the second sample date.

The backward selection process resulted in the following variables being identified as significant predictors ($p < 0.1$) of Bray-Curtis dissimilarity: Secchi depth, water temperature, change in water temperature, dissolved oxygen concentration, change in dissolved oxygen concentration, dissolved oxygen saturation, change in dissolved oxygen saturation, $NO_3 + NO_2$ concentration, change in $NO_3 + NO_2$ concentration, $NH_4$ concentration, change in $NH_4$

concentration, total phosphorus concentration, change in total phosphorus concentration, change

in dissolved reactive phosphorus, the interaction between water temperature and Secchi depth,

and the interaction between water temperature and dissolved oxygen concentration. The adjusted

$R^2$ value of this model was 0.229. See table below for associated significance values.

|  | Estimate | Std. Error | t value | p value | |
|---|---|---|---|---|---|
| (Intercept) | 0.130915 | 0.584248 | 0.224 | 0.823026 | |
| secchi | -0.057016 | 0.01867 | -3.054 | 0.002705 | ** |
| wtemp | 0.014433 | 0.026594 | 0.543 | 0.588196 | |
| wtemp.change | 0.057769 | 0.019769 | 2.922 | 0.004054 | ** |
| o2 | -0.825094 | 0.383033 | -2.154 | 0.032944 | * |
| o2.change | 0.192872 | 0.0795 | 2.426 | 0.016536 | * |
| o2sat | 0.121762 | 0.050084 | 2.431 | 0.016314 | * |
| o2sat.change | -0.016267 | 0.006772 | -2.402 | 0.017611 | * |
| no3no2.change | -0.245014 | 0.137362 | -1.784 | 0.076638 | . |
| no3no2 | -0.306475 | 0.11621 | -2.637 | 0.009303 | ** |
| nh4 | -0.625507 | 0.305741 | -2.046 | 0.042641 | * |
| nh4.change | 0.726915 | 0.346381 | 2.099 | 0.037648 | * |
| totp | 2.537219 | 1.011803 | 2.508 | 0.013297 | * |
| totp.change | 2.69869 | 1.125084 | 2.399 | 0.017772 | * |
| drp.change | -2.409117 | 1.223509 | -1.969 | 0.050925 | . |
| wtemp:secchi | 0.00439 | 0.001242 | 3.534 | 0.000554 | *** |
| wtemp:o2 | -0.031129 | 0.010366 | -3.003 | 0.003169 | ** |

legend: $p < 0.1$:  .
$p < 0.05$:  *
$p < 0.01$:  **
$p < 0.001$: ***

**Appendix 6: Lake Mendota 16S rRNA Gene Sequencing Time Series Analysis**

Sample Collection

We used the Lake Mendota bacterial 16S rRNA gene sequencing time series (96 samples with > 7000 taxa over 11 years) for this analysis. Microbial cells were collected on 0.2 um filters from 12-m integrated water column samples as previously described (Kara et al 2013).

Data Processing

Sample processing, sequencing and core amplicon data analysis were performed by the Earth Microbiome Project (EMP) (www.earthmicrobiome.org) (Gilbert et al 2014), and all amplicon sequence data and metadata have been made public through the data portal (qiita.microbio.me/emp). Briefly, community DNA was used to amplify partial 16S rRNA genes using the 515F-806R primer pair (Caporaso et al 2011) and an Illumina MiSeq, with standard EMP protocols. Processed data were obtained from the data portal (study ID 1242) and included 45,094,125 total and 3,058,149 unique sequences. These were grouped into 7,081 OTUs using de novo picking in QIIME version 2.0.6 with the deblur workflow. Three samples with fewer than 100 reads were removed (06 October 2006, 13 July 2006, and 04 December 2009) prior to deblurring.

Analysis

We removed two additional failed samples (08 July 2004 and 10 May 2008) that had fewer than 100 reads after deblurring. All other samples had > 10,000 reads. Despite different sequencing depths across samples, we did not rarefy the dataset. We suggest that the cohesion

workflow be used with non-rarefied datasets, because rarefying introduces additional bias into pairwise correlations (see readme).

Before beginning the connectedness calculations, we removed all taxa with a mean abundance lower than $10^{-7}$. This removed an average of 17.5% of the relative abundance of each sample. We repeated the analysis with mean abundance thresholds of $10^{-8}$, $10^{-8.5}$, and $10^{-9}$, and we found qualitatively consistent results among these different thresholds. We set the persistence cutoff for inclusion of taxa at presence in at least 5 samples. However, due to the high sequencing depth of this dataset, only two additional OTUs were removed due to the persistence cutoff. We note that the values of Bray-Curtis dissimilarity used in the regression were calculated on the full dataset prior to removing any taxa due to these cutoffs.

We included samples in the Bray-Curtis dissimilarity regression if another sample was taken within 25 to 41 days. Samples were taken approximately monthly during the ice-free season, so this range encompassed the majority of samples (54 of 91). We calculated cohesion using both the taxon shuffle null model and the row shuffle null model. Both options showed that cohesion was a significant predictor of Bray-Curtis dissimilarity; for both null models, weaker positive cohesion and stronger negative cohesion significantly related to lower Bray-Curtis dissimilarity. In both cases, negative cohesion was the much stronger predictor, although both variables were significant. The model fit was slightly better for the row shuffle null model (adjusted $R^2 = 0.35$ for taxon shuffle and adjusted $R^2 = 0.38$ for row shuffle), so we present the analyses using the row shuffle null model in the main text.

References:

Caporaso, J. Gregory, et al. "Global patterns of 16S rRNA diversity at a depth of millions of
sequences per sample." *Proceedings of the National Academy of Sciences*
108.Supplement 1 (2011): 4516-4522.

Gilbert, Jack A., Janet K. Jansson, and Rob Knight. "The Earth Microbiome project: successes
and aspirations." *BMC biology* 12.1 (2014): 69.

Kara, Emily L., et al. "A decade of seasonal dynamics and co-occurrences within freshwater
bacterioplankton communities from eutrophic Lake Mendota, WI, USA." *The ISME
journal* 7.3 (2013): 680-684.

**Chapter 5:** Small subsets of highly connected taxa predict compositional change in microbial communities

**Authors:** Cristina M. Herren[1], Katherine D. McMahon[2]


**Author affiliations:**

[1]Freshwater and Marine Sciences Program, University of Wisconsin - Madison, Madison, Wisconsin, USA

[2]Departments of Bacteriology and Civil and Environmental Engineering, University of Wisconsin - Madison, Madison, Wisconsin, USA

**Abstract**

For decades, ecological theory has predicted that the complexity of communities should be related to their stability. However, this prediction has rarely been tested empirically, because of both the difficulty of finding suitable systems where the question is tractable and the trouble of defining "stability" in real systems. Microbial communities provide the opportunity to investigate a related question: how does community connectivity relate to the rate of compositional turnover? We used a newly developed metric called community "cohesion" to test how microbial community connectivity relates to Bray-Curtis dissimilarity through time. In three long-term datasets, we found that stronger connectivity corresponded to lower rates of compositional turnover. Using two case studies of disturbed and reference communities, we found that the predictive power of community connectivity was diminished by external disturbance. Finally, we tested whether the highly connected taxa were disproportionately important in explaining compositional turnover. We found that subsets of highly connected "keystone" taxa, generally comprising 1-5% of community richness, explained community turnover better than using all taxa. Our results suggest that stronger biotic interactions within microbial community dynamics are stabilizing to community composition, and that highly connected taxa are good indicators of pending community shifts.

**Introduction**

Theoretical ecologists have studied the relationship between community complexity and stability for decades (MacArthur 1955, May 1972, Pimm 1979, Neutel et al. 2007). Initial results suggesting that complex communities should be unstable (May 1972) prompted a rich literature aimed at understanding how complex communities persist in nature. The primary source of "complexity" considered in these studies is the strength of species interactions (May 2001). These theoretical studies consistently find that connectivity arising from species interactions is a major contributor to community stability (McCann et al. 1998, Ives et al. 2000, Neutel et al. 2002, Wiliams and Martinez 2004). Depending on the configuration and strength of species interactions within a community, greater connectivity can lead to increased or diminished stability (Allesina and Tang 2012).

Despite the substantial theoretical literature on how complexity influences stability, comparatively few studies have investigated this question empirically (but see Kondoh 2008, Neutel and Thorne 2014, Jacquet et al. 2016). This is partly due to the logistical challenges of addressing this question in real systems; such challenges include the difficulty in quantifying species interactions (Laska and Wootton 1998), the need to observe many taxa to satisfy model assumptions (Allesina and Tang 2012), the difficulty of sampling communities completely (Polis 1991), and the need to collect data spanning many generations of the study organisms (Morin and Lawler 1995). Another practical hurdle is defining the terms "complexity" and "stability" for real communities (Connell and Sousa 1983, Neubert and Caswell 1997). Studies that have tested how community complexity relates to community stability have found mixed results. Recently, an analysis of 116 food webs found no consistent pattern between complexity and stability (Jacquet et al. 2016). However, prior studies have found evidence of positive (Polis and Strong

1996, Fagan 1997, Dunne et al. 2002) and negative (Pimm and Lawton 1978, Stouffer and

Bascompte 2011) relationships. Thus, relatively little of the ecological theory regarding the

complexity-stability debate has been tested empirically, and results of these empirical studies are

mixed.

Microbial communities are promising systems for investigating the relationship between

community structure and community stability. Several characteristics of microbial communities

make it possible to overcome the previously described challenges of testing theoretical

hypotheses in empirical systems; microbial communities are sufficiently diverse as to meet the

richness assumptions of theoretical models, hundreds of generations can be observed within one

dataset, and the resolution of next generation sequencing datasets means that even rare taxa (<

0.01% of communities) are sampled. However, one prominent challenge of testing ecological

hypotheses in microbial communities is that interactions between taxa are difficult to observe

and therefore must be inferred from observed population dynamics. For this reason, we

previously created a robust metric, called "cohesion," that quantifies the instantaneous

connectivity of microbial communities (Herren and McMahon *in press*).  Briefly, this method

quantifies connectedness values for each taxon in a dataset based on its average correlations with

other taxa. Cohesion metrics are calculated from the abundance and connectedness of the taxa

present in each community. When many highly connected taxa are present, the cohesion values

for a community are larger in magnitude. There are two cohesion values for each sample,

corresponding to connectivity arising from positive taxon relationships and connectivity arising

from negative taxon relationships.

Recent studies have hypothesized that biotic interactions are important for mediating

compositional stability in microbial communities. For example, Zelezniak et al. (2015) found

that persistent sub-networks within microbial communities often included a high degree of
facilitation among taxa. This result suggested that facilitation reinforces existing community
composition, leading to lower rates of compositional change. Furthermore, several microbial
studies have found evidence of "keystone taxa," which are highly interactive and have a
disproportionate effect on their communities (Vick-Majors et al. 2014, Agler et al. 2016,
Banerjee et al. 2016). Changes in the abundance of keystone taxa lead to shifts in community
composition due to cascading effects on other taxa (Mills et al. 1993). Finally, viruses and
protists constitute a major source of mortality in marine bacterial communities (Fuhrman and
Noble 1995, Suttle 2007), indicating the importance of predation in shaping community
composition. Thus, multiple lines of evidence suggest that the strength of biotic interactions
within microbial communities should be related to the rate of compositional change.

In this study, we use three long-term microbial datasets (each spanning 10+ years) to test
the hypothesis that higher connectivity in microbial communities is related to greater
compositional stability through time. As mentioned previously, it is difficult to quantify stability
for empirical systems. Instead, we use a related metric as our response variable, which is
compositional turnover through time (Bray-Curtis dissimilarity). Modeling Bray-Curtis
dissimilarity is a major aim of microbial ecology (Larsen et al. 2012), because the function of
microbial communities is expected to change in parallel with changes in community composition
(Urich et al. 2008, Sekirov et al. 2010). We also asked whether highly connected "keystone" taxa
are disproportionally important for explaining compositional turnover (Power et al. 1996, Jordán
et al. 1999). To answer this question, we repeated our analyses of community connectivity versus
compositional stability using only the highly connected taxa. Finally, we reasoned that the
influence of biotic interactions on compositional change would be diminished when external

disturbance to a community was high (Dai et al. 2017). To test this hypothesis, we analyzed two case studies where communities experienced different levels of disturbances. We hypothesized that connectivity would be a better predictor of compositional change when external forcing, and disturbance, was lower. Together, these analyses aimed to identify the conditions under which connectivity is related to compositional turnover and to investigate which taxa are most informative about overall community changes.

**Methods**

*Datasets*

To test our hypotheses about 1) the relationship between connectivity and compositional turnover and 2) the influence of highly connected taxa, we obtained three long-term, publicly available microbial datasets. These included the San Pedro Ocean Time Series bacterial dataset (SPOT) from the coastal ocean near southern California (described in detail in Cram et al. 2015a), the Lake Mendota (Wisconsin, USA) phytoplankton dataset (ME-phyto, described in detail at https://lter.limnology.wisc.edu), and the Lake Mendota bacterial dataset (ME-bact, described in detail at https://lter.limnology.wisc.edu). Additional information and references for datasets can be found in the Supplementary Online Materials (SOM). We chose these datasets because of their long duration (SPOT: 10 years, ME-phyto: 19 years, ME-bact: 11 years), their large number of samples (SPOT: 274 samples with 437 taxa, ME-phyto: 293 samples with 409 taxa, ME-bact: 91 samples with 7081 taxa), and the variety of technologies used to obtain the datasets (SPOT: automated ribosomal intergenic spacer analysis [ARISA], ME-phyto: cell counts under microscope, ME-bact: 16S rRNA gene amplicon sequencing).

We identified two case studies where comparable microbial communities experienced differing levels of external disturbance. The first case study is the comparison of the phytoplankton communities in Peter Lake and Paul Lake in northern Wisconsin, USA (described in Elser and Carpenter 1988, Cottingham et al. 1998). These lakes were originally one water body, but were artificially divided into two lakes for the purpose of conducting ecological disturbance experiments. Paul Lake served as the undisturbed reference system, while Peter Lake was experimentally disturbed using nutrient supplementation and fish additions over the course of the time series (see SOM). Each lake was sampled 197 times over 12 years. Phytoplankton taxa were enumerated using direct cell counts under a microscope.

The second disturbance case study is a comparison between two types of plaque communities sampled as a part of the Human Microbiome Project (HMP). Briefly, samples were collected from 242 human volunteers at up to 18 body sites at two sample collection dates with a maximum interval of 14 days. We compared the bacterial communities from the highly disturbed, exposed plaque site (supragingival plaque) to the protected plaque site beneath the gums (subgingival plaque). For both sites, we evaluated the relationship between community cohesion and compositional turnover (Bray-Curtis dissimilarity) in an individual's microbiome between the two sampling times.

*Hypotheses: Long-Term Datasets*

Following the result that persistent microbial sub-networks are enriched in taxon interactions (Zelezniak et al. 2015), we expected that greater connectivity would be related to lower compositional change. Additionally, we hypothesized that the highly connected taxa would have a disproportionate influence on community dynamics. Thus, we expected that

subsets of highly connected taxa would be better predictors of community turnover (Bray-Curtis dissimilarity) than randomly chosen subsets of taxa.

*Hypotheses: Case Study Comparisons*

For Peter Lake and Paul Lake, we reasoned that the experimental perturbations would be a cause of community composition change in Peter Lake, but not in the undisturbed Paul Lake. Therefore, we expected that biotic interactions would contribute less to compositional turnover in the disturbed lake, Peter Lake. Thus, we hypothesized that community cohesion would be a better predictor of Bray-Curtis dissimilarity in the undisturbed Paul Lake than in Peter Lake.

For the two plaque bacterial communities, we reasoned that compositional change at the exposed site (supragingival plaque) would be influenced more strongly by immigration and dispersal than by biotic interactions. Conversely, we expected that the protected plaque communities (subgingval plaque) would be influenced by biotic interactions, because taxa are contained in close proximity for long periods of time. Thus, we hypothesized that cohesion would be a significant predictor of Bray-Curtis dissimilarity for the protected plaque site (subgingival plaque), but not at the exposed site (supragingival plaque).

*Statistical Methods*

We used cohesion metrics (Herren and McMahon *in press*) as a measure of the connectivity of the microbial communities (see SOM). We calculated cohesion metrics for the five datasets (three long-term time series and two case studies). Briefly, this workflow calculates two metrics for each sample quantifying the connectivity due to positive correlations between taxa and connectivity due to negative correlations between taxa. Cohesion metrics are calculated for each sample by taking the sum of every taxon's connectedness score (also calculated within the cohesion workflow) multiplied by its abundance in the sample.

For each dataset, we conducted linear regressions modeling the compositional turnover (Bray-Curtis dissimilarity) between time points as a function of the cohesion metrics. Stated another way, we asked whether cohesion metrics predict Bray-Curtis dissimilarity. For the SPOT dataset, we analyzed the bacterial communities from the chlorophyll maximum site, reasoning that the cholorophyll maximum site represented a discrete ecological community. For taxa in the HMP plaque datasets, we calculated taxon connectedness values using correlations between taxa among individuals at the first sampling timepoint. Additional methods and the parameter values used in the workflow for each dataset can be found in Supplementary Table 1.

To test the hypothesis that highly connected taxa are disproportionately influential in determining community dynamics, we iteratively repeated the regression analysis (modeling Bray-Curtis dissimilarity as a function of community cohesion), each time calculating cohesion from different subsets of taxa. We excluded taxa based on their connectedness values, where we removed the least connected taxa first. For example, when 40 taxa were included in the analysis, the negative cohesion metric was calculated from the 40 taxa with the strongest negative connectedness, and the positive cohesion metric was calculated from the 40 taxa with the strongest positive connectedness. We recorded the $R^2$ value from the linear model (Bray-Curtis dissimilarity vs. cohesion) for each subset of taxa.

We then repeated the workflow described above (removing taxa and running the linear regression) using random subsets of taxa, rather than using the most highly connected taxa. Thus, when 40 taxa were included, we randomly selected 40 taxa from which to calculate the positive and negative cohesion values. We recorded the model $R^2$ value of the linear regression when taxa were randomly included in the workflow. Then, we repeated this process 500 times, as to generate a distribution of model $R^2$ values when 40 random taxa were selected. We ran 500

models for each possible number of taxa included in the workflow. We had hypothesized that the highly connected taxa would be more informative about overall community changes than randomly chosen taxa; thus, we expected that the model using the highly connected taxa would have a larger $R^2$ value.

**Results**

*Long-Term Datasets*

For each of the three long-term datasets (SPOT, ME-phyto, and ME-bact), we used linear regression to analyze the amount of variability in community composition turnover (Bray-Curtis Dissimilarity) that could be explained by community connectivity (cohesion metrics). Representative results from all datasets analyzed are presented in Table 1.

Cohesion was a significant predictor of Bray-Curtis dissimilarity in all three long-term datasets. Stronger cohesion, whether positive or negative, was consistently related to lower rates of compositional change (Table 1). Stronger negative cohesion was significantly related to lower Bray-Curtis dissimilarity in all three datasets (Fig. 1B, D, F). In the ME-bact dataset, stronger positive cohesion was also significantly related to lower compositional turnover (Table 1). Maximum adjusted model $R^2$ values were 0.485 for ME-phyto, 0.428 for ME-bact, and 0.478 for SPOT chlorophyll maximum.

We re-calculated cohesion metrics from subsets of highly connected taxa in order to evaluate whether highly connected taxa were disproportionately informative about compositional turnover (black line, Fig. 1A, C, E). We also calculated cohesion metrics using random subsets of taxa to evaluate whether highly connected taxa modeled Bray-Curtis dissimilarity better than randomly chosen taxa (grey lines, Fig. 1A, C, E). In the models containing random subsets of

taxa, model $R^2$ values declined as fewer taxa were included in cohesion calculations (solid grey line indicates the median). Conversely, in models using the most highly connected taxa, the adjusted $R^2$ values remained stable as the least-connected taxa were removed (black line). In all three long-term datasets, adjusted $R^2$ values increased when small subsets (< 5% total richness) of highly connected taxa were included (Table 1). Maximum $R^2$ values occurred when using 15 taxa in ME-phyto, 33 taxa in ME-bact, and 15 taxa in SPOT (Fig. 1A, C, E).

In all three datasets, models based on the most highly connected taxa to calculate cohesion significantly outperformed the models using random subsets of taxa when small proportions of taxa were included. Significance was determined as instances when the model $R^2$ value using highly connected taxa was above the 95th percentile of $R^2$ values from models using random taxa. For the SPOT dataset, the model using highly connected taxa performed significantly better than the model using randomly selected taxa when fewer than 25 taxa were included. For the ME-phyto dataset, it was when fewer than 35 taxa were included. For the ME-bact dataset, it was fewer than 105 taxa.

*Identities of Highly Connected Taxa*

We were curious about the identities of the most highly connected taxa in the three long-term datasets. We focused on taxa that had the strongest negative associations with other taxa, because negative cohesion was highly significant in all long-term datasets (Fig. 1, Table 1). In the ME-phyto dataset, eight of the ten taxa with the largest negative connectedness values were cyanobacteria (see SOM for list). For the ME-bact dataset, we compared the lists of the fifty most abundant taxa and the fifty taxa with largest negative connectedness values (see SOM). Twenty-two taxa were on both these lists. Twenty-eight taxa were among the fifty most

connected but not the fifty most abundant. These included three of the four recognized clades in the acIV Actinobacteria lineage, a member of the Chloroflexi phylum, and two members of the Planctomycetes phylum, all of which are relatively understudied by freshwater microbial ecologists. Among the Proteobacteria in this list were PnecD, a relatively rare member of the genus *Polynucleobacter*, and several members of the order Rhizobiales. Although these organisms are not among the most ubiquitous or abundant taxa found in freshwater lakes, the results obtained here motivate us to study their ecology more intently, particularly with genome-based methods.

*Case Study: Peter Lake and Paul Lake*

As with the long-term datasets, we used cohesion metrics as predictors of Bray-Curtis dissimilarity for phytoplankton communities in Peter and Paul Lakes. We had hypothesized that cohesion metrics would be better predictors of compositional change in the reference system, Paul Lake. We conducted separate analyses for the two lakes.

As expected, cohesion metrics were better predictors for Paul Lake than for the disturbed system, Peter Lake. We evaluated this prediction by comparing model $R^2$ values for the two lakes (Fig. 2). Across nearly the entire range of taxa included, models analyzing the Bray-Curtis dissimilarity of phytoplankton communities in Paul Lake had a higher $R^2$ value than similar models for Peter Lake. The exception was when very few (< 10) taxa were included in the cohesion calculations. In both lakes, model $R^2$ values dropped significantly when fewer than 10 taxa were used to calculate cohesion. The best model fit in Paul Lake occurred when 13 taxa were included (adjusted $R^2 = 0.487$), whereas for Peter Lake it was 57 taxa (adjusted $R^2 = 0.374$). In Paul Lake, stronger negative cohesion and weaker positive cohesion was both

significantly related to lower Bray-Curtis dissimilarity (Table 1). In Peter Lake, stronger negative cohesion and stronger positive cohesion were both significantly related to lower Bray-Curtis dissimilarity (Table 1).

*Case Study: Exposed and Protected Plaque Communities*

We tested whether cohesion could explain community composition turnover in plaque communities in the human-associated microbiome. We expected that cohesion would be a significant predictor of compositional turnover at the protected plaque site (subgingival plaque), but not at the exposed plaque site (supragingival plaque). In this analysis, we calculated Bray-Curtis dissimilarities from two communities sampled from the same individual host, collected at two different time points.

In the exposed plaque communities (supragingival plaque), we found that there was no significant relationship between either cohesion metric and Bray-Curtis dissimilarity (Fig. 3). However, in the protected plaque communities (subgingival plaque), cohesion was significantly related to Bray-Curtis dissimilarity. The model fit was best (adjusted $R^2$ = 0.207) when 13 OTUs were included (Fig. 3). Stronger positive cohesion and weaker negative cohesion were both significantly related to lower Bray-Curtis dissimilarity (Table 1).

We conducted this same analysis using the other 16 body sites sampled as a part of the Human Microbiome Project (see SOM for results). Most sites (11 of 16) showed highly significant relationships ($p < 0.001$) between cohesion and the rate of compositional turnover (Bray-Curtis dissimilarity). At all 11 sites, stronger negative cohesion was related to lower Bray-Curtis dissimilarity. Positive cohesion was highly significant at 6 of the 11 sites, but showed mixed relationships with Bray-Curtis dissimilarity.

**Discussion**

  The consistent results from the three long-term (10+ year) microbial time series showed that stronger connectivity within aquatic microbial communities was related to greater compositional stability. In all three cases, stronger cohesion values were significantly related to lower Bray-Curtis dissimilarity over time (Fig. 1B, D, F). Moreover, models using information from small subsets of highly connected taxa predicted compositional turnover performed better than models using all taxa (Fig. 1A, C, E). Therefore, the most highly connected taxa had the strongest relationship with compositional change, and their presence corresponded to increased compositional stability. In all three long-term datasets, highly connected taxa performed significantly better than models built using random assemblages of taxa. Only a small fraction of taxa, generally comprising 1-5% of total richness, were necessary to model compositional turnover. These qualitatively consistent results show support for the hypotheses that 1) community connectivity is a strong mediator of compositional stability and 2) highly connected taxa have disproportionate influence on observed community dynamics.

  The predictive power of our models in the long-term datasets was striking, given that no environmental factors were included in these analyses. For the three long-term datasets, the model $R^2$ values ranged between 0.4 and 0.5. For comparison, previous analyses modeling the community similarity between time points in the SPOT dataset obtained maximum $R^2$ values of approximately 0.2, even when using over 30 environmental parameters (Cram et al. 2015a). Similarly, a model explaining compositional turnover in the ME-phyto dataset using environmental variables had an adjusted $R^2$ value of 0.23 (Herren and McMahon *in press*).

Our result that stronger negative cohesion was related to lower compositional turnover in the long-term time series was consistent across a variety of ecosystems, sampling methods, and sample dates. The three datasets were obtained using different techniques for determining abundance, including direct cell counts (ME-phyto), 16S rRNA gene tag sequencing (ME-bact), and ARISA (SPOT). These methods all differ in their sensitivity and bias. Thus, the consistency of our results suggest that including cohesion as a predictor variable might improve models of compositional turnover in many microbial systems.

*Disturbance Decreases the Importance of Connectivity*

The case studies of disturbed systems showed that community cohesion had less explanatory power when communities experienced external disturbance. The Peter Lake vs. Paul Lake comparison showed that cohesion metrics were better predictors of Bray-Curtis dissimilarity in the undisturbed system, Paul Lake (Fig. 2). In Peter Lake, experimental perturbations caused shifts in the phytoplankton community (Carpenter et al. 1987, Carpenter et al. 1996, Cottingham and Carpenter 1998). Thus, some of the compositional change in Peter Lake was due to experimental disturbances. Our results agree with these previous conclusions, suggesting that connectivity had decreased influence on compositional change in the perturbed lake, Peter Lake.

Analyses of the protected and exposed plaque sites showed that community cohesion was only an important explanatory factor in compositional turnover at the protected plaque site (Fig. 3). Many of the same OTUs were present in the protected and exposed plaque communities, but their connectedness and power to predict compositional change were different at the two sites.

These results suggest that high levels of disturbance and dispersal can disrupt the relationship between biotic interactions and community stability.

There are two main ways in which disturbance can alter the relationship between biotic interactions and compositional change. First, disturbances causing high immigration or emigration of taxa disrupt established species interactions. Biotic interactions drive population dynamics by influencing taxon growth and death rates (Gotelli 2001); thus, the effects of biotic interactions will be most apparent when taxa interact consistently over many generations. Second, disturbances cause compositional change that is not linked to biotic interactions. For example, compositional change at the exposed plaque site may have resulted from tooth brushing or from consuming food. Thus, the proportion of total compositional change due to biotic interactions would be diminished in this case. The lower predictive power of cohesion when applied to highly disturbed communities suggests that the importance of biotic interactions in community assembly and turnover is context dependent.

*Highly Connected Taxa as Keystone Taxa*

By calculating cohesion using different subsets of taxa, we identified the taxa that contributed most to the relationship between connectivity and turnover. In all three long-term datasets, the maximum model fit occurred when a small number (15-33) of taxa were included. Similarly, in the two reference systems in the case study analyses, the optimal number of taxa to include was 13 for both datasets (Table 1). Thus, focusing on these highly connected taxa may be a useful strategy for researchers seeking to understand microbial community assembly. Including taxa with lower connectedness values in our models often obscured the signal of connectivity captured by the cohesion metrics. These results support the hypothesis that these highly

connected taxa may function as "keystone taxa" within microbial communities; the relatively small subsets of highly connected taxa had outsized explanatory power of overall community dynamics. Additionally, although some of the datasets contained the same phytoplankton taxa (ME-phyto, Peter Lake, Paul Lake), the same taxon received different scores of connectedness in the various datasets. This result suggests that the ecological context of the microbial communities is important for determining which taxa will act as keystone taxa in various environments.

Using fewer taxa in the cohesion calculations often improved the fit of our models. We propose that this approach of evaluating model fit using different subsets of taxa could be generalized to other analyses with different response variables. Model fit should be best when the most informative taxa are selected. One strategy for identifying taxa with disproportionate influence would be to include the taxa where the model $R^2$ values spike in Fig. 1. Although model $R^2$ values remained high when small numbers of taxa were included, we would caution against building predictive models with fewer than 5-10 taxa. In this case, cohesion values obtained from a training set of communities may be prone to high variability when applied to new communities, especially if there are directional trends in taxon abundances over time.

*Ecological Interpretation of Connectivity and Compositional Turnover*

In the majority of instances where cohesion metrics were significant predictors of Bray-Curtis dissimilarity, stronger connectivity was related to greater compositional stability. However, there were cases that deviated from this norm, where stronger connectivity was related to more rapid change. We hypothesize that these anomalies are mediated by the ecology of the different study sites. For example, the result from Paul Lake that stronger positive cohesion was

destabilizing was driven by samples from the summer of 1993, when a large and persistent cyanobacterial bloom disrupted normal seasonal dynamics. Similarly, following the result that cohesion had lower explanatory power in disturbed systems, it would be interesting to investigate how the strength of deterministic versus stochastic forces alters the relationship between community connectivity and community stability. This might be done with the Human Microbiome Project dataset, as immigration and selective pressure likely differ between body sites (Li and Ma 2016). Our preliminary analysis of this dataset showed that 12 of the 18 sites had a strong relationship between cohesion values and compositional turnover rate, but the explanatory power of the models varied. Quantifying dispersal and selection rates at different sites may shed light on the variability of the observed relationships and the degree to which community connectivity can explain compositional change.

Under the assumption that cohesion measures biotic interactions (Herren and McMahon *in press*), our results support the hypothesis that biotic interactions are stabilizing to microbial community composition. Several recent studies have concluded that biotic interactions can be strong drivers of microbial population dynamics, on par with or exceeding the influence of environmental factors (Cram et al. 2015b, Lima-Mendez et al. 2015, Weitz et al. 2016, Cabello et al. 2016, Trivedi et al. 2017). For example, many OTUs are more strongly related to other OTUs than to habitat variables (Cram et al. 2015b). However, few studies have tested the relationship between connectivity and compositional change, primarily because the methods to quantify connectivity have only been recently developed. Initial theoretical studies indicated that stronger biotic linkages would be destabilizing to ecological communities (May 1972, Pimm 1979). However, these initial studies also made several simplifying assumptions about the organization of ecological food webs. The ensuing literature has discovered several possible

mechanisms that allow diverse and complex communities to persist in nature (e.g. McCann et al. 1998, Brose et al. 2006, Kondoh 2006). Future work might consider how the attributes of microbial communities, including spatial structuring (Long and Azam 2001), dispersal rates (Finlay 2002), and the possibility of dormancy (Lennon and Jones 2011) influence the relationship between connectivity and compositional stability.

Biotic interactions create feedback loops within ecological communities that can amplify or dampen the effects of external perturbations (Berryman and Millstein 1989). One mechanistic hypothesis for the result that stronger connectivity is related to lower compositional change is that the taxon interactions in microbial communities are arranged to form negative feedback loops, thereby mitigating the effects of disturbance (Konopka et al. 2015, Coyte et al. 2015). Thus, stronger interactions would lead to stronger negative feedback loops that buffer communities from compositional change. Our findings also agree with recent work showing that persistent modules of taxa are enriched in taxon interactions (Zelezniak et al. 2015). Thus, another interpretation is that biotic interactions create self-reinforcing modules within bacterial communities, which leads to lower turnover. One possible mechanism generating these self-reinforcing subunits is metabolite exchange between taxa (Morris et al. 2013, Levy and Borenstein 2013). Finally, our work agrees with recent studies that hypothesize that microbial communities contain keystone taxa, which shape community assembly due to their strong interactions with other taxa (Vick-Majors et al. 2014, Agler et al. 2016, Banerjee et al. 2016). We propose that studying these keystone taxa might allow researchers to prioritize organism-centric studies to learn why and how specific taxa have such a strong influence on communities.

Ecological theory offers some insight into why negative cohesion was often more strongly related to compositional stability than positive cohesion. Under some circumstances,

pairwise correlations may be indicative of pairwise taxon interactions; we make this simplifying assumption to investigate our results in the context of classical ecological theory. Mathematical models using local stability analysis with simple communities have indicated that stable equilibria are common when negative interactions (e.g. competition, predation) are present. For instance, scenarios with stable eqiulibria include: two or more competitors (May and Leonard 1975), one predator and one prey (Rosenzweig and MacArthur 1963), one predator with multiple prey (Holt 1977), and multiple predators with one or more prey (McPeek 2012). Conversely, stability is rare in food webs with exclusively positive pairwise interactions (May 1981). However, recent theoretical literature has indicated that mutualism within the context of other, negative interactions can be stabilizing (Mougi and Kondoh 2012). Thus, ecological theory indicates that the placement and strength of negative interactions within communities is critical to maintaining stable composition. The traits of the most highly connected Mendota phytoplankton taxa further support this line of reasoning. Most of the taxa associated with low compositional turnover were cyanobacteria, which often have a competitive relationship with other phytoplankton (Fong et al. 1993). Several studies have documented the self-reinforcing effect of competition for light in aquatic environments, showing that high cyanobacterial abundance can be a stable state in eutrophic lakes (Scheffer et al. 1997, Schröder et al. 2005). Thus, our results align with existing theoretical explanations of phytoplankton community transitions, and suggest that similar dynamics may be present in other systems. Although the knowledge of traits and interactions is scarce for taxa in the other two long-term datasets (SPOT and Mendota-bact), the lists of highly connected taxa provided here (available in SOM) may be useful starting places for trait-based studies.

We encourage future studies to examine traits of highly connected taxa using modeling or experimental approaches. Strong correlations between taxa are often construed as interactions between taxa. Although this assumption can be useful for invoking ecological theory, there are several conditions where this assumption would be false. For example, two competing taxa might show a negative correlation in their abundances through time due to competitive exclusion; conversely, two competing taxa may have similar niches, and therefore might show a positive correlation due to simultaneous responses to environmental drivers. Finally, other trophic levels likely influence the correlations and connectedness metrics observed in these microbial communities, although these factors are not explicitly included in these analyses. Thus, mechanistic models would greatly benefit further studies of the role of highly connected taxa in community dynamics.

Our results show several empirical instances where stronger connectivity is related to greater compositional stability, contrary to the initial theoretical finding that highly connected communities should be unstable. Empirical food webs have many non-random attributes, which may explain why our results differ from theoretical expectations and analyses of simulated datasets (Pimm 1980, Polis 1991, Neutel et al. 2007). Our consistent finding that greater community connectivity (especially from negative connections between taxa) results in lower compositional turnover suggests that either evolutionary or community assembly processes arrange biotic interactions to form stabilizing feedback loops.

**Acknowledgements**

Supplementary material is available at ISME Journal's website.

**References**

Agler MT, Ruhe J, Kroll S, Morhenn C, Kim S-T, Weigel D, *et al*. (2016). Microbial Hub Taxa Link Host and Abiotic Factors to Plant Microbiome Variation. *PLOS Biol* **14**: e1002352.

Allesina S, Tang S. (2012). Stability criteria for complex ecosystems. *Nature* **483**: 205–208.

Banerjee S, Kirkby CA, Schmutter D, Bissett A, Kirkegaard JA, Richardson AE. (2016). Network analysis reveals functional redundancy and keystone taxa amongst bacterial and fungal communities during organic matter decomposition in an arable soil. *Soil Biology and Biochemistry* **97**: 188–198.

Berryman AA, Millstein JA. (1989). Are ecological systems chaotic — And if not, why not? *Trends in Ecology & Evolution* **4**: 26–28.

Brose U, Williams RJ, Martinez ND. (2006). Allometric scaling enhances stability in complex food webs. *Ecology Letters* **9**: 1228–1236.

Cabello AM, Cornejo-Castillo FM, Raho N, Blasco D, Vidal M, Audic S, *et al*. (2016). Global distribution and vertical patterns of a prymnesiophyte–cyanobacteria obligate symbiosis. *ISME J* **10**: 693–706.

Carpenter SR, Kitchell JF, Cottingham KL, Schindler DE, Christense DL, Post DM, *et al*. (1996). Chlorophyll Variability, Nutrient Input, and Grazing: Evidence from Whole-Lake Experiments. *Ecology* **77**: 725–735.

Carpenter SR, Kitchell JF, Hodgson JR, Cochran PA, Elser JJ, Elser MM, *et al*. (1987). Regulation of Lake Primary Productivity by Food Web Structure. *Ecology* **68**: 1863–1876.

Connell JH, Sousa WP. (1983). On the Evidence Needed to Judge Ecological Stability or Persistence. *The American Naturalist* **121**: 789–824.

Cottingham KL, Carpenter SR. (1998). Population, Community, and Ecosystem Variates as Ecological Indicators: Phytoplankton Responses to Whole-Lake Enrichment. *Ecological Applications* **8**: 508–530.

Cottingham KL, Carpenter SR, Amand ALS. (1998). Responses of epilimnetic phytoplankton to experimental nutrient enrichment in three small seepage lakes. *J Plankton Res* **20**: 1889–1914.

Coyte KZ, Schluter J, Foster KR. (2015). The ecology of the microbiome: Networks, competition, and stability. *Science* **350**: 663–666.

Cram JA, Chow C-ET, Sachdeva R, Needham DM, Parada AE, Steele JA, *et al*. (2015a). Seasonal and interannual variability of the marine bacterioplankton community throughout the water column over ten years. *ISME J* **9**: 563–580.

Cram JA, Xia LC, Needham DM, Sachdeva R, Sun F, Fuhrman JA. (2015b). Cross-depth analysis of marine bacterial networks suggests downward propagation of temporal changes. *ISME J* **9**: 2573–2586.

Dai W, Zhang J, Tu Q, Deng Y, Qiu Q, Xiong J. (2017). Bacterioplankton assembly and interspecies interaction indicating increasing coastal eutrophication. *Chemosphere* **177**: 317–325.

Dunne JA, Williams RJ, Martinez ND. (2002). Network structure and biodiversity loss in food webs: robustness increases with connectance. *Ecology Letters* **5**: 558–567.

Elser JJ, Carpenter SR. Predation-driven dynamics of zooplankton and phytoplankton communities in a whole-lake experiment. *Oecologia* **76**: 148–154.

Fagan WF. (1997). Omnivory as a Stabilizing Feature of Natural Communities. *The American Naturalist* **150**: 554–567.

Finlay BJ. (2002). Global Dispersal of Free-Living Microbial Eukaryote Species. *Science* **296**: 1061–1063.

Fong P, Donohoe RM, Zedler JB. (1993). Competition with macroalgae and benthic cyanobacterial mats limits phytoplankton abundance in experimental microcosms. *Marine Ecology Progress Series* **100**: 97–102.

Fuhrman JA, Noble RT. (1995). Viruses and protists cause similar bacterial mortality in coastal seawater. *Limnol Oceanogr* **40**: 1236–1242.

Gotelli NJ. (1995). A primer of ecology. Sinauer Associates Incorporated. Sunderland, MA, USA.

Herren CM, McMahon KD. (2017). Cohesion: A method for quantifying the connectivity of microbial communities. *ISME J in press*

Holt RD. (1977). Predation, apparent competition, and the structure of prey communities. *Theoretical Population Biology* **12**: 197–229.

Ives AR, Klug JL, Gross K. (2000). Stability and species richness in complex communities. *Ecology Letters* **3**: 399–411.

Jacquet C, Moritz C, Morissette L, Legagneux P, Massol F, Archambault P, *et al*. (2016). No complexity-stability relationship in empirical ecosystems. *Nat Commun* **7**: 12573.

Jordán F, Takács-Sánta A, Molnár I. (1999). A Reliability Theoretical Quest for Keystones. *Oikos* **86**: 453–462.

Kondoh M. (2008). Building trophic modules into a persistent food web. *PNAS* **105**: 16631–16635.

Kondoh M. (2006). Does foraging adaptation create the positive complexity–stability relationship in realistic food-web structure? *Journal of Theoretical Biology* **238**: 646–651.

Konopka A, Lindemann S, Fredrickson J. (2015). Dynamics in microbial communities: unraveling mechanisms to identify principles. *ISME J* **9**: 1488–1495.

Larsen PE, Field D, Gilbert JA. (2012). Predicting bacterial community assemblages using an artificial neural network approach. *Nat Meth* **9**: 621–625.

Laska MS, Wootton JT. (1998). Theoretical Concepts and Empirical Approaches to Measuring Interaction Strength. *Ecology* **79**: 461–476.

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

Levy R, Borenstein E. (2013). Metabolic modeling of species interaction in the human microbiome elucidates community-level assembly rules. *PNAS* **110**: 12804–12809.

Li L, Ma Z (Sam). (2016). Testing the Neutral Theory of Biodiversity with Human Microbiome Datasets. *Scientific Reports* **6**: 31448.

Lima-Mendez G, Faust K, Henry N, Decelle J, Colin S, Carcillo F, *et al*. (2015). Determinants of community structure in the global plankton interactome. *Science* **348**: 1262073.

Long RA, Azam F. (2001). Antagonistic Interactions among Marine Pelagic Bacteria. *Appl Environ Microbiol* **67**: 4975–4983.

MacArthur R. (1955). Fluctuations of Animal Populations and a Measure of Community Stability. *Ecology* **36**: 533–536.

May R, Leonard W. (1975). Nonlinear Aspects of Competition Between Three Species. *SIAM J Appl Math* **29**: 243–253.

May RM. (1981). Theoretical Ecology. Sinauer. Sunderland, MA, USA.

May RM. (2001). Stability and Complexity in Model Ecosystems. Princeton University Press. Princeton, NJ, USA.

May RM. (1972). Will a Large Complex System be Stable? *Nature* **238**: 413–414.

McCann K, Hastings A, Huxel GR. (1998). Weak trophic interactions and the balance of nature. *Nature* **395**: 794–798.

McPeek MA. (2012). Intraspecific density dependence and a guild of consumers coexisting on one resource. *Ecology* **93**: 2728–2735.

Mills LS, Soulé ME, Doak DF. (1993). The Keystone-Species Concept in Ecology and Conservation. *BioScience* **43**: 219–224.

Morin PJ, Lawler SP. (1995). Food Web Architecture and Population Dynamics: Theory and Empirical Evidence. *Annual Review of Ecology and Systematics* **26**: 505–529.

Morris BEL, Henneberger R, Huber H, Moissl-Eichinger C. (2013). Microbial syntrophy: interaction for the common good. *FEMS Microbiology Reviews* **37**: 384–406.

Mougi A, Kondoh M. (2012). Diversity of Interaction Types and Ecological Community Stability. *Science* **337**: 349–351.

Neubert MG, Caswell H. (1997). Alternatives to Resilience for Measuring the Responses of Ecological Systems to Perturbations. *Ecology* **78**: 653–665.

Neutel A-M, Heesterbeek JAP, van de Koppel J, Hoenderboom G, Vos A, Kaldeway C, *et al.* (2007). Reconciling complexity with stability in naturally assembling food webs. *Nature* **449**: 599–602.

Neutel A-M, Heesterbeek JAP, Ruiter PC de. (2002). Stability in Real Food Webs: Weak Links in Long Loops. *Science* **296**: 1120–1123.

Neutel A-M, Thorne MAS. (2014). Interaction strengths in balanced carbon cycles and the absence of a relation between ecosystem complexity and stability. *Ecol Lett* **17**: 651–661.

Pimm SL. (1979). Complexity and Stability: Another Look at MacArthur's Original Hypothesis. *Oikos* **33**: 351–357.

Pimm SL. (1980). Properties of Food Webs. *Ecology* **61**: 219–225.

Pimm SL, Lawton JH. (1978). On feeding on more than one trophic level. *Nature* **275**: 542–544.

Polis GA. (1991). Complex Trophic Interactions in Deserts: An Empirical Critique of Food-Web Theory. *The American Naturalist* **138**: 123–155.

Polis GA, Strong DR. (1996). Food Web Complexity and Community Dynamics. *The American Naturalist* **147**: 813–846.

Power ME, Tilman D, Estes JA, Menge BA, Bond WJ, Mills LS, *et al*. (1996). Challenges in the Quest for Keystones. *BioScience* **46**: 609–620.

Rosenzweig ML, MacArthur RH. (1963). Graphical Representation and Stability Conditions of Predator-Prey Interactions. *The American Naturalist* **97**: 209–223.

Scheffer M, Rinaldi S, Gragnani A, Mur LR, van Nes EH. (1997). On the Dominance of Filamentous Cyanobacteria in Shallow, Turbid Lakes. *Ecology* **78**: 272–282.

Schröder A, Persson L, De Roos AM. (2005). Direct experimental evidence for alternative stable states: a review. *Oikos* **110**: 3–19.

Sekirov I, Russell SL, Antunes LCM, Finlay BB. (2010). Gut Microbiota in Health and Disease. *Physiological Reviews* **90**: 859–904.

Stouffer DB, Bascompte J. (2011). Compartmentalization increases food-web persistence. *PNAS* **108**: 3648–3652.

Suttle CA. (2007). Marine viruses — major players in the global ecosystem. *Nat Rev Micro* **5**: 801–812.

Trivedi P, Delgado-Baquerizo M, Trivedi C, Hamonts K, Anderson IC, Singh BK. (2017). Keystone microbial taxa regulate the invasion of a fungal pathogen in agro-ecosystems. *Soil Biology and Biochemistry* **111**: 10–14.

Urich T, Lanzén A, Qi J, Huson DH, Schleper C, Schuster SC. (2008). Simultaneous Assessment of Soil Microbial Community Structure and Function through Analysis of the Meta-Transcriptome. *PLOS ONE* **3**: e2527.

Vick-Majors TJ, Priscu JC, Amaral-Zettler LA. (2014). Modular community structure suggests metabolic plasticity during the transition to polar night in ice-covered Antarctic lakes. *ISME J* **8**: 778–789.

Weitz JS, Stock CA, Wilhelm SW, Bourouiba L, Coleman ML, Buchan A, *et al*. (2015). A multitrophic model to quantify the effects of marine viruses on microbial food webs and ecosystem processes. *ISME J* **9**: 1352–1364.

Williams RJ, Martinez ND, Morin AEPJ. (2004). Limits to Trophic Levels and Omnivory in Complex Food Webs: Theory and Data. *The American Naturalist* **163**: 458–468.

Zelezniak A, Andrejev S, Ponomarova O, Mende DR, Bork P, Patil KR. (2015). Metabolic dependencies drive species co-occurrence in diverse microbial communities. *PNAS* 201421834.

**Figure 1:** Analyses of the three long-term microbial datasets show that stronger cohesion is related to lower compositional turnover in all three long-term datasets. Left-hand panels (A, C, E) show the how the adjusted model $R^2$ values of the regression analysis changed as taxa were excluded from cohesion calculations. For each number of taxa on the x-axis, cohesion values were calculated from the most highly connected taxa (black line) and from a random subset of taxa (grey lines). The solid grey line shows the median adjusted model $R^2$ for randomly selected subsets, while the dashed grey lines give the 5% and 95% intervals. Median $R^2$ values from models using random subsets of taxa declined as fewer taxa were included in the cohesion metrics. When 1-5% of taxa within a community were used to calculate cohesion, models using highly connected taxa generally had higher model $R^2$ values than models using random taxa. The red stars in left-hand panels identify the regression model with the highest adjusted $R^2$, which is displayed in the paired right-hand panel. Right-hand panels (A, C, E) show the best-fitting linear regressions modeling compositional turnover (Bray-Curtis dissimilarity) as a function of cohesion from negative connections between taxa. Points indicate Bray-Curtis dissimilarity between sequential samples. Solid lines show the fit of linear models. All three datasets showed that cohesion arising from negative correlations between taxa was a strong predictor of Bray-Curtis dissimilarity (Table 1).

**Figure 2:** Cohesion explained a greater amount of variability in phytoplankton community turnover in the undisturbed Paul Lake, as compared to an experimentally disturbed system, Peter Lake. The model $R^2$ values predicting Bray-Curtis dissimilarity in Paul Lake were generally higher than for models predicting Bray-Curtis dissimilarity in Peter Lake. The exception was when models used very few (< 10) taxa to calculate cohesion metrics. As in Figure 1, taxa were

sequentially removed from the analysis in reverse order of their connectedness (i.e. least connected taxon removed first).

**Figure 3:** The adjusted model $R^2$ values for plaque communities sampled as part of the Human Microbiome Project show that cohesion was a significant predictor of Bray-Curtis dissimilarity in the protected plaque site (subgingival plaque, solid line), but not at the exposed plaque site (supragingival plaque, dashed line). Icons above the solid line indicate when positive cohesion was significant at $p < 0.001$ (+) and when negative cohesion was significant at $p < 0.001$ (-). At the exposed plaque site, cohesion was never a significant ($p < 0.05$) predictor of Bray-Curtis dissimilarity.

**Figure 1:**
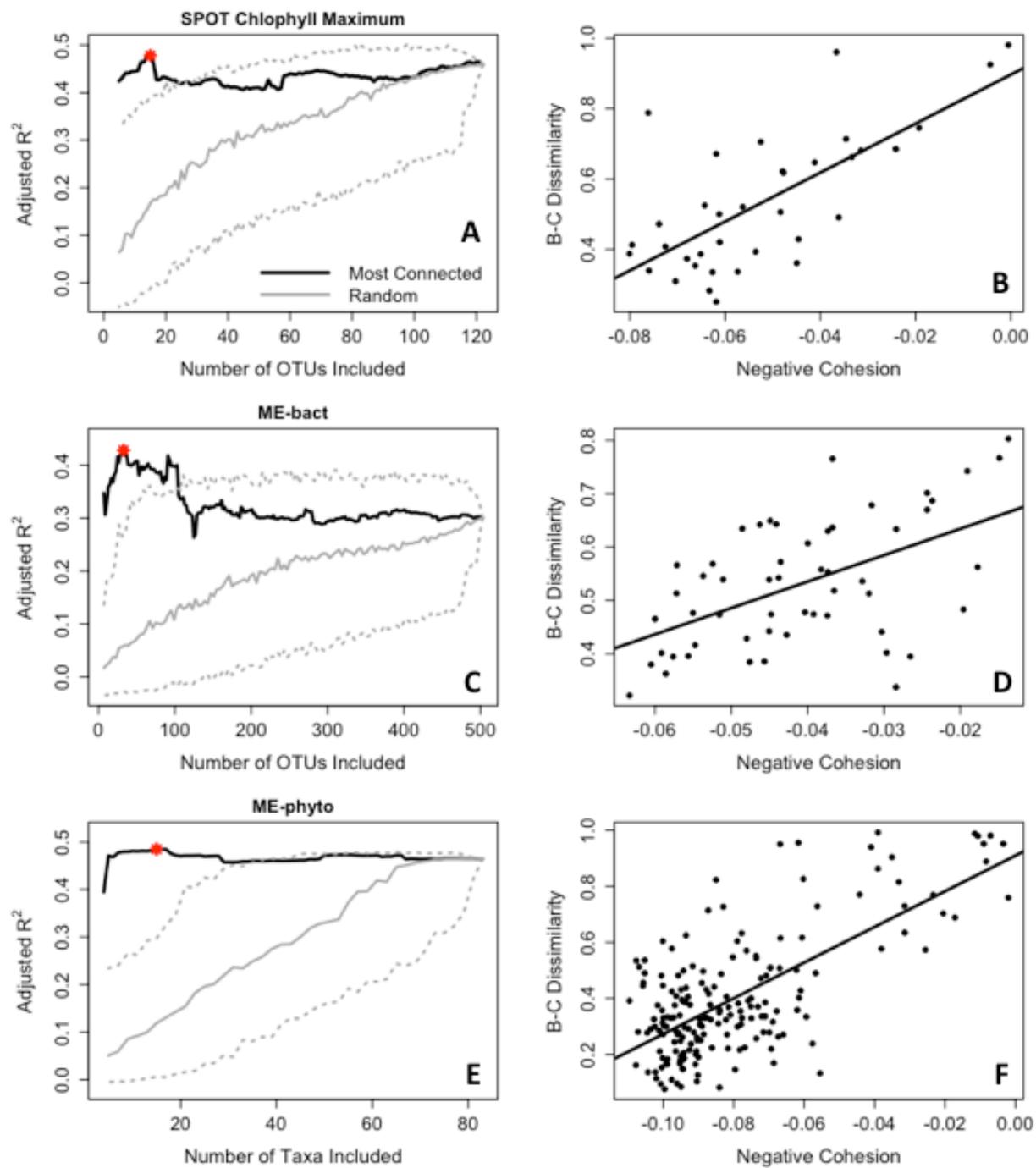
**Figure 2:**
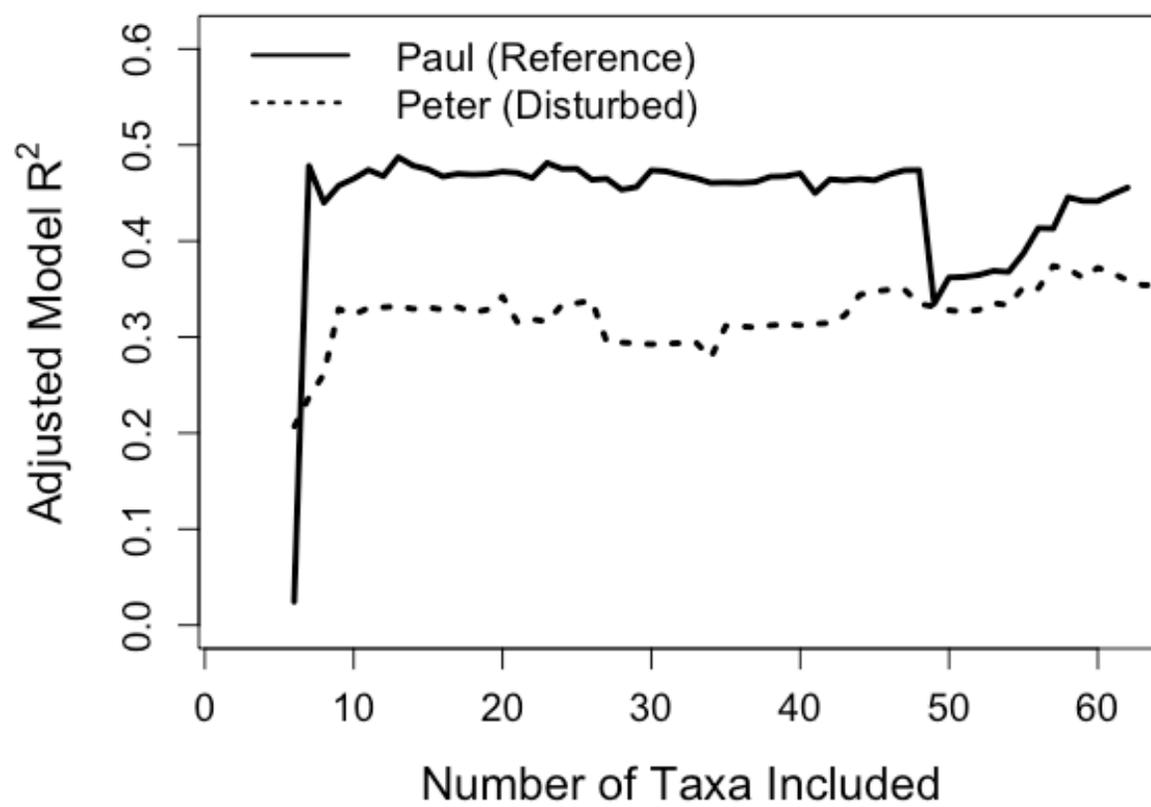
**Figure 3:**
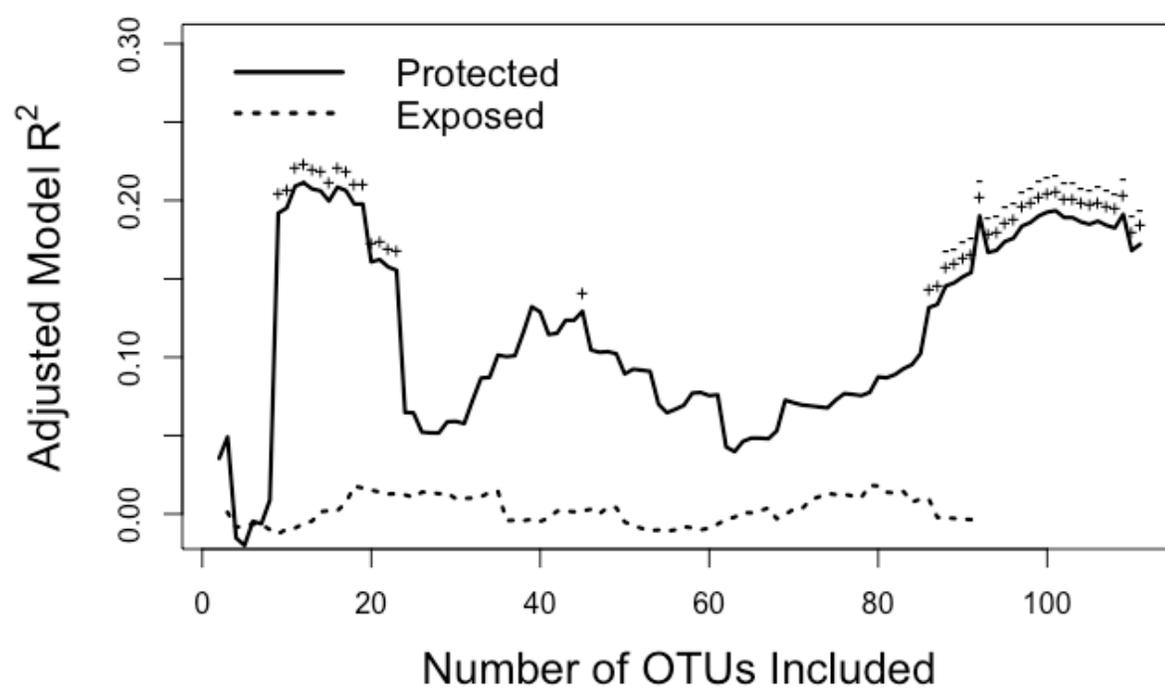
**Table 1: Representative Results of Cohesion as a Predictor of Bray-Curtis Dissimilarity**

| Dataset | Total # of Taxa | Optimal # of Taxa* | Maximum Adjusted $R^2$ | Positive Cohesion P value | Negative Cohesion P value | Positive Cohesion Direction [+] | Negative Cohesion Direction [+] | Data Points in Analysis [#] |
|---|---|---|---|---|---|---|---|---|
| ME - phyto | 409 | 15 | 0.485 | n.s. | $< 1*10^{-27}$ | NA | Stronger is stabilizing | 186 |
| ME - bact | 7081 | 33 | 0.428 | $< 1*10^{-3}$ | $< 1*10^{-7}$ | Stronger is stabilizing | Stronger is stabilizing | 54 |
| SPOT – Chl. Max. | 392 | 15 | 0.478 | n.s. | $< 1*10^{-4}$ | NA | Stronger is stabilizing | 36 |
| Protected Plaque | 2190 | 13 | 0.207 | $< 1*10^{-4}$ | 0.014 | Stronger is stabilizing | Weaker is stabilizing | 93 |
| Exposed Plaque | 2124 | 79 | 0.018 | n.s. | n.s. | NA | NA | 95 |
| Paul - Reference | 209 | 13 | 0.487 | $< 1*10^{-5}$ | $< 1*10^{-18}$ | Weaker is stabilizing | Stronger is stabilizing | 123 |
| Peter - Disturbed | 237 | 57 | 0.374 | 0.009 | $< 1*10^{-6}$ | Stronger is stabilizing | Stronger is stabilizing | 121 |

\* Indicates the number of taxa where the maximum adjusted $R^2$ value occurred

+ These columns indicate the direction of a significant relationship between cohesion and Bray-Curtis dissimilarity. For example, "stronger is stabilizing" means that greater cohesion is related to lower Bray-Curtis dissimilarity. Non-significant relationships are denoted "n.s.".

# Because the time elapsed between samples is strongly related to Bray-Curtis dissimilarity, we only included paired sampled with similar time separation (see Supplementary Table 1 for further information). This reduced the number of data points available for our analyses.

**Appendix 1: Supplementary Table 1: Data Processing and Cohesion Workflow Parameters**

| Dataset | Presence Cutoff | Abundance Cutoff | # Taxa (Total) | # Taxa Meeting Cutoffs | # Samples for Correlations | Range of Days Elapsed | Null Model Algorithm |
|---|---|---|---|---|---|---|---|
| ME - phyto | 15 | None | 409 | 83 | 293 | 36 – 48 | Column Shuffle |
| ME - bact | 5 | $1* 10^{-9}$ | 7081 | 503 | 91 | 25 – 41 | Row Shuffle |
| SPOT – Chl. Max. | 5 | $1* 10^{-3}$ | 392 | 122 | 65 | 20 – 35 | Row Shuffle |
| Protected Plaque | 18 | $1* 10^{-8}$ | 2190 | 111 | 171 | max. of 14 | Row Shuffle |
| Exposed Plaque | 18 | $1* 10^{-8}$ | 2124 | 91 | 174 | max. of 14 | Row Shuffle |
| Paul - Reference | 18 | None | 209 | 62 | 196 | 39 - 45 | Column Shuffle |
| Peter - Disturbed | 18 | None | 237 | 65 | 195 | 39 - 45 | Column Shuffle |

Column Descriptions

*Presence Cutoff:* The minimum number of samples in which a taxon must be present in order to be included in the cohesion calculation

*Abundance Cutoff:* The minimum mean abundance for a taxon across all samples in order for the taxon to be included in the cohesion calculation

*# Taxa (Total):* The total number of taxa observed in the dataset

*# Taxa Meeting Cutoffs:* The number of taxa that satisfied the presence cutoff and the abundance cutoff, and therefore were included in the cohesion calculations

*# Samples for Correlations:* The number of samples used to calculate pairwise taxon correlations

*Range of Days Elapsed:* Time between samples affects the observed dissimilarity between samples. Thus, we only calculated Bray-Curtis dissimilarity for samples within a given time frame. The time frames were based primarily on sampling frequency.

*Null Model Algorithm:* There are two possible randomization algorithms implemented in null model within the cohesion workflow. The benefits of each algorithm are outlined in the readme

file in the SOM of Herren and McMahon 2017. This variable indicates whether the column

randomization or row randomization was used in the null model.

**Appendix 2: Supplementary Material and Methods**

*Dataset Descriptions and References*

Lake Mendota Phytoplankton:

We obtained these data from the North Temperate Lakes Long Term Ecological Research (NTL-LTER) data portal, https://lter.limnology.wisc.edu/data. Associated protocols can also be located there. Briefly, whole water samples were collected from a central location in Lake Mendota. The top 0-8m of the water column was collected, and a sample was preserved for analysis by PhycoTech Inc. Phytoplankton counts were obtained by mounting samples on slides and enumerating the cells using a microscope. Samples were taken approximately every 2 weeks during the open water (ice-free) season spanning the years 1995 – 2013. We removed counts not identified at any level (e.g. categorized as "Miscellaneous"). We did not alter or combine identifications from those presented in the dataset. For each sample, we transformed phytoplankton counts to relative abundances by dividing the count for each taxon by the total number of counts for the entire sample.

San Pedro Ocean Time Series:

We obtained these data from the following data repository: http://www.bco-dmo.org/dataset/535915. Additional protocols and information about the San Pedro Ocean time series can be found at the following webpage, hosted by the University of Southern California: https://dornsife.usc.edu/spot/data/. Further description of bacterial community dynamics and environmental conditions can be found in Fuhrman et al. 2006, Chow et al. 2013, and Cram et al. 2015. Briefly, water samples from five depths were taken at the same location in the San Pedro ocean channel (off the coast of southern California) from the years 2000 - 2011. Bacterial

community composition was determined using automated ribosomal interspacer analysis (ARISA). We decided to analyze the samples from the chlorophyll maximum (CMAX) because we hypothesized that this site would represent a coherent community; the other sites were sampled at constant depths. There were 65 samples at the CMAX site that spanned the years 2000 – 2009.

Lake Mendota 16S rRNA gene sequencing:

We obtained these data from the Earth Microbiome Project data portal, study ID 1242. During quality control processing, we removed five samples with fewer than 100 reads. Details on sample collection and data processing can be found in the Supplementary Online Material of Herren and McMahon, *in press.* Taxonomy assignment was conducted using a custom workflow available at https://github.com/McMahonLab/TaxAss.

Peter Lake Phytoplankton and Paul Lake Phytoplankton:

We obtained these data from the NTL-LTER data portal, but data were generated by the Cascade research group (https://cascade.limnology.wisc.edu/). Phytoplankton counting protocols are detailed in Cottingham et al. 1998.

Briefly, three water samples were drawn from different depths in the epilimnion at a central station in each lake. Water samples were pooled and preserved for later identification. Phytoplankton were filtered from the water and mounted on slides to be enumerated under a microscope. Samples were collected approximately weekly during the summer stratified season (May – September) from the years 1984 – 1995. For each sample, we transformed phytoplankton

counts to relative abundances by dividing the count for each taxon by the total number of counts for the entire sample.

Peter Lake was disturbed using a number of experimental perturbations over the course of the time series. These disturbances targeted multiple aspects of the lake food web. Initial disturbances focused on the trophic cascade induced by changing the top predators in an ecosystem (Carpenter et al. 1985). In 1985, researchers removed most piscivorous fish from Peter Lake, replacing them with zooplanktivorous fish (Carpenter et al. 1987). Then, in the springs of 1988 and 1989, trout were added to Peter Lake. Manipulations of the fish community continued through 1990, including the addition and removal of various species (Carpenter and Kitchell 1996). Later manipulations evaluated the influence of bottom-up dynamics by introducing additional nutrients. In 1993 and 1994, Peter Lake received daily additions of nitrogen and phosphorus (Cottingham et al. 1998).

Human Microbiome Plaque 16S rRNA gene sequencing:

We obtained the HMP data from the associated website, http://hmpdacc.org. The version of the data downloaded was the 16S variable region 3-5 (V35) clustered with mothur. Information about sample collection, processing, and data generation can also be found at http://hmpdacc.org. We removed pilot samples from the dataset before conducting our analyses. We removed all samples lacking a second time point. For each body site, we calculated correlations between taxa using samples from the first of the two time points. The lowest number of reads in a sample of these two plaque communities was 783, and all other samples had >1000 reads. Thus, we did not remove any samples from our analyses due to low sequencing depth.

References:

Carpenter SR, Kitchell JF. (1996). The Trophic Cascade in Lakes. Cambridge University Press.

Carpenter SR, Kitchell JF, Hodgson JR. (1985). Cascading Trophic Interactions and Lake Productivity. *BioScience* **35**: 634–639.

Carpenter SR, Kitchell JF, Hodgson JR, Cochran PA, Elser JJ, Elser MM, *et al*. (1987). Regulation of Lake Primary Productivity by Food Web Structure. *Ecology* **68**: 1863–1876.

Chow CET, Sachdeva R, Cram JA, Steele JA, Needham DM, 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.

Cottingham KL, Carpenter SR, Amand ALS. (1998). Responses of epilimnetic phytoplankton to experimental nutrient enrichment in three small seepage lakes. *J Plankton Res* **20**: 1889–1914.

Cram JA, Chow CET, Sachdeva R, Needham DM, Parada AE, Steele JA, *et al*. (2015). Seasonal and interannual variability of the marine bacterioplankton community throughout the water column over ten years. *ISME J* **9**: 563–580.

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

Herren CM, McMahon KD. (2017). Cohesion: A method for quantifying the connectivity of microbial communities. *ISME J in press*

**Appendix 3: Human Microbiome Project Supplementary Analysis**

We applied the same workflow as described for the Supragingival Plaque and Subgingival Plaque communities to the other 16 body sites sampled as a part of the Human Microbiome Project. Briefly, we calculated cohesion values using subsets of highly connected taxa at each body site. We recorded the adjusted model $R^2$ values for the multiple regressions predicting Bray-Curtis dissimilarity within a single host as a function of positive and negative cohesion. The parameters used in the cohesion workflow were the same as for the plaque communities.

Here, we report the body sites for which positive and negative cohesion were highly significant ($P < 0.001$) predictors for any of the regression models. To evaluate the influence of the null model used to correct pairwise correlations, we conducted this analysis using either 1) the row shuffle null model or 2) the column shuffle null model. We found that the choice of null model had little influence on the results of these analyses; the body sites where positive and negative cohesion were significant were identical when using the two different null models. Icons above the plotted lines in the top panels of supplementary figures indicate when positive (+) or negative (-) cohesion was significant at $P < 0.001$. Plotted lines in the bottom panels of supplementary figures show slope estimates for the cohesion variables for different numbers of included taxa. When the slope estimate for negative cohesion was positive, stronger negative cohesion was related to lower compositional change. Conversely, when the slope estimate for positive cohesion was negative, stronger positive cohesion was related to lower compositional change.

Positive Cohesion Significant (6): Hard palate, Left antecubital fossa, Right retroauricular crease, Saliva, Stool, Throat

Negative Cohesion Significant (11): Hard palate, Left antecubital fossa, Left retroauricular crease, Mid vagina, Posterior fornix, Right antecubital fossa, Right retroauricular crease, Saliva, Stool, Throat, Vaginal Introitus

## Appendix 4: Highly Connected Taxa

Mendota Phytoplankton Dataset:

|  | Negative Connectedness |
|---|---|
| Chroococcaceae | -0.116147747 |
| Synechococcus.sp1 | -0.100280376 |
| Microcystis.aeruginosa.1 | -0.068239572 |
| Rhodomonas.minuta.var..nannoplanctica | -0.067932942 |
| Aphanothece.1 | -0.05249947 |
| Cryptomonas.erosa.1 | -0.048318822 |
| Aphanothece.nidulans.1 | -0.043411415 |
| Phormidium.mucicola | -0.041645258 |
| Synechocystis | -0.041562774 |
| Synechocystis.1 | -0.039788644 |
| Rhodomonas.minuta | -0.034053018 |
| Aphanocapsa.delicatissima.1 | -0.033701429 |
| Chlamydomonas.1 | -0.032959032 |
| Chlamydomonas.globosa | -0.030972652 |
| Oscillatoria.limnetica.1 | -0.030632259 |
| Cryptomonas.erosa | -0.030095405 |
| Monoraphidium.capricornutum | -0.028807359 |
| Cryptomonas.rostratiformis.1 | -0.026791256 |
| Aphanizomenon.flos.aquae | -0.026421031 |

Mendota 16S Bacterial Dataset:

| V3 | V4 | V5 | V6 | V7 | V8 |
|---|---|---|---|---|---|
| p__Actinobacteria | c__Actinobacteria | o__Actinomycetales(100) | acI(100) | acI-B(100) | acI-B1(100) |
| p__Proteobacteria | c__Alphaproteobacteria | o__Rickettsiales(100) | alfV(100) | alfV-A(100) | LD12(100) |
| p__Actinobacteria | c__Actinobacteria | o__Actinomycetales(100) | acI(100) | acI-A(99) | Phila(97) |
| p__Actinobacteria | c__Actinobacteria | o__Actinomycetales(100) | acTH1(100) | acTH1-A(100) | acTH1-A1(100) |
| p__Bacteroidetes | c__Saprospirae | o__[Saprospirales](100) | bacI(100) | bacI-A(100) | bacI-A1(99) |
| p__Proteobacteria | c__Betaproteobacteria | o__Burkholderiales(100) | betII(96) | Pnec(96) | PnecC(74) |
| p__Actinobacteria | c__Acidimicrobiia | o__Acidimicrobiales(100) | acIV(100) | acIV-C(100) | Iluma-C1(99) |
| p__Actinobacteria | c__Actinobacteria | o__Actinomycetales(100) | acI(100) | acI-A(100) | acI-A6(100) |
| p__Bacteroidetes | c__Saprospirae | o__[Saprospirales](100) | bacI(100) | bacI-A(100) | bacI-A1(99) |
| p__Actinobacteria | c__Actinobacteria | o__Actinomycetales(100) | acI(100) | acI-A(98) | acI-A4(97) |
| p__Actinobacteria | c__Actinobacteria | o__Actinomycetales(100) | acI(100) | acI-A(100) | acI-A5(97) |
| p__Proteobacteria | c__Betaproteobacteria | o__Burkholderiales(100) | betI(100) | betI-A(100) | Lhab-A3(54) |
| p__Actinobacteria | c__Actinobacteria | o__Actinomycetales(100) | acI(100) | acI-A(100) | acI-A7(59) |
| p__Proteobacteria | c__Betaproteobacteria | o__Methylophilales(96) | betIV(96) | betIV-A(96) | LD28(96) |
| p__Proteobacteria | c__Betaproteobacteria | o__Burkholderiales(96) | betIII(90) | betIII-A(90) | betIII-A1(87) |
| p__Actinobacteria | c__Actinobacteria | o__Actinomycetales(100) | acI(100) | acI-A(90) | acI-A3(77) |
| p__Proteobacteria | c__Betaproteobacteria | o__Burkholderiales(100) | betI(100) | betI-A(100) | Lhab-A1(89) |
| p__Proteobacteria | c__Betaproteobacteria | o__Burkholderiales(100) | betIII(100) | betIII-A(100) | betIII-A1(98) |
| p__Bacteroidetes | c__Cytophagia | o__Cytophagales(100) | bacIII(100) | bacIII-A(100) | |
| p__Actinobacteria | c__Acidimicrobiia | o__Acidimicrobiales(100) | acIV(100) | acIV-B(94) | Iluma-B1(93) |

SPOT Chlorophyll Maximum:

| nodeIDs | Phylum | Class | Order |
|---|---|---|---|
| CMAX_ARISA_686.9 | Proteobacteria | Alphaproteobacteria | Rickettsiales |
| CMAX_ARISA_666.4 | Proteobacteria | Alphaproteobacteria | Rickettsiales |
| CMAX_ARISA_435.5 | Actinobacteria | Actinobacteria | koll13 |
| CMAX_ARISA_682.4 | Proteobacteria | Alphaproteobacteria | Rickettsiales |
| CMAX_ARISA_662 | Proteobacteria | Alphaproteobacteria | Rickettsiales |
| CMAX_ARISA_538.9 | Proteobacteria | Gammaproteobacteria | Oceanospirillales |
| CMAX_ARISA_570.1 | Cyanobacteria | Chloroplast | Chlorophyta |
| CMAX_ARISA_674.2 | Proteobacteria | Alphaproteobacteria | Rickettsiales |
| CMAX_ARISA_671.2 | NA | NA | NA |
| CMAX_ARISA_424.4 | Actinobacteria | Actinobacteria | koll13 |
| CMAX_ARISA_721.2 | Proteobacteria | Gammaproteobacteria | Chromatiales |
| CMAX_ARISA_421.5 | Actinobacteria | Actinobacteria | koll13 |
| CMAX_ARISA_679.4 | Proteobacteria | Alphaproteobacteria | Rickettsiales |
| CMAX_ARISA_561.8 | Cyanobacteria | Chloroplast | Chlorophyta |
| CMAX_ARISA_478.8 | Proteobacteria | Deltaproteobacteria | Bdellovibrionales |
| CMAX_ARISA_741.8 | Proteobacteria | Deltaproteobacteria | Desulfobacterales |
| CMAX_ARISA_670.5 | Proteobacteria | Alphaproteobacteria | Rickettsiales |
| CMAX_ARISA_653.1 | SAR406 | AB16 | ZA3648c |
| CMAX_ARISA_729.4 | Proteobacteria | Deltaproteobacteria | Desulfobacterales |
| CMAX_ARISA_423.3 | Actinobacteria | Actinobacteria | koll13 |

**Appendix 5: R Script Used in Keystone Taxa Analysis**

#This script shows our analysis of the Lake Mendota phytoplankton dataset.

#It also generates the figure showing the Model R2 value predicting Bray-Curtis dissimilarity

using different numbers of taxa.

# CMH 23-June-2017


```r
library(reshape)
library(lme4)
library(vegan)

## create necessary functions ##
#This function counts the number of zeroes in a vector
zero <- function(vec){
  num.zero <- length(which(vec == 0))
  return(num.zero)
}

#This function averages only negative values in a vector
neg.mean <- function(vector){
  neg.vals <- vector[which(vector < 0)]
  n.mean <- mean(neg.vals)
  if(length(neg.vals) == 0) {n.mean <- 0}
  return(n.mean)
}

#create function that averages only positive values in a vector
pos.mean <- function(vector){
  pos.vals <- vector[which(vector > 0)]
  p.mean <- mean(pos.vals)
  if(length(pos.vals) == 0) {p.mean <- 0}
  return(p.mean)
}

#This function pulls out a quantile from each vector in a matrix
prop.vec <- function(mat, prop){
  prop.number <- round(prop*dim(mat)[1])
  sort.mat <- apply(mat, 2, sort)
  prop.vector <- sort.mat[prop.number, ]
  return(prop.vector)
}
```

```
## Parameter inputs ##
#decide whether to do row or col shuffle
col.shuffle <- T
#pick a proportion of presence to analyze taxon
zero.prop <- .05
#Set days elapsed options for paired samples (for calculating Bray-Curtis dissimilarity)
elapse.min <- 36
elapse.max <- 48


## Read in phyto data
p <-
        read.csv("~/Desktop/UWMadison/McMahonLab/Var_and_R2/MethodsManuscript/Thesi
        sRevisions/FilesForUpload/SOM/OnlineCSV/MendotaPhytoTable.csv", header = T,
        row.names = 1)


#Convert to matrix
b <- as.matrix(p)
#Take out blank rows and columns
b <- b[rowSums(b) > 0, colSums(b) > 0]
#order by date
b <- b[order(as.Date(rownames(b), format = "%m/%d/%Y")), ]


#save a copy of original dataset before further manipulation
orig <- b
orig.rel <- orig / rowSums(orig)


#define persistence cutoff for retaining a taxon in the analysis
zero.cutoff <- ceiling(zero.prop * dim(b)[1])


#cut out low persistence taxa
c <- b[ , apply(b, 2, zero) < (dim(b)[1]-zero.cutoff) ]
#Take out any rows that are now blank
c <- c[rowSums(c) > 0, ]


#convert to relative abundance
rel.c <- c / rowSums(orig)
#optionally, check to see what proportion of community was removed
#hist(rowSums(rel.c))


#create observed - expected correlations in relative data to compare to res.cor
d.sub <- as.matrix(rel.c)


#create relative abundance correlation matrix
cor.mat.true <- cor(d.sub)


otu.cors.true.pos <- apply(cor.mat.true, 1, pos.mean)
```

```
otu.cors.true.neg <- apply(cor.mat.true, 1, neg.mean)

#define number of permutations for null model
perm.cor <- 200

#create vector to hold median otu-otu correlations for each otu
med.otu.cors <- vector()

#create random number vector to hold seeds for randomization
#set seed here to create same seeds in each run
set.seed(888)
seeds <- sample(seq(1:100000000), size = (dim(d.sub)[2] * perm.cor * dim(d.sub)[1]))

if(!col.shuffle){
 #This is the row-shuffle null model
 for(m in 1:dim(d.sub)[2]){
   which.otu <- m

   #create vector to hold correlations from every permutation for each single otu
   perm.cor.vec.mat <- vector()

   #Run this loop as many times as specified by perm.cor
   for(i in 1:perm.cor){
     #Duplicate the d.sub matrix, and then replace entries with randomized numbers
     d.sub2 <- d.sub
     for(j in 1:dim(d.sub)[1]){
       #Randomize only taxa present in sample (i.e. abundance > 0)
       which.replace <- which(d.sub2[j, ] > 0 )
       #Do not randomize abundance of focal taxon
       if(d.sub2[j, m] > 0) {which.replace <- which.replace[!which.replace == m]}
       which.replace.minus <- which.replace[!(which.replace %in% m)]
       #Set new seed for upcoming randomization
       set.seed(seeds[j + (i-1)*(dim(d.sub)[1]) + (m-1)*dim(d.sub)[2] ])
       #Replace original values with randomized values
       d.sub2[j, which.replace.minus ] <- sample(d.sub[ j, which.replace.minus])

     }
     #replace focal column with original column, just to be sure it stays the same
     d.sub2[, which.otu] <- d.sub[ , which.otu]

     #calculate correlations between randomized taxon vectors
     cor.mat.true.null <- cor(d.sub2)

     #save the vector corresponding to the focal taxon, m
     perm.cor.vec.mat <- cbind(perm.cor.vec.mat, cor.mat.true.null[,m])
```

```
    }
    #Take the median of the correlations produced by the null model
    med.otu.cors <- cbind(med.otu.cors, apply(perm.cor.vec.mat, 1, median))

    if(m %% 20 == 0){print(m)}
  }

} else {
  #This is the column shuffle null model
  for(m in 1:dim(d.sub)[2]){
    which.otu <- m

    #create vector to hold correlations from every permutation for each single otu
    perm.cor.vec.mat <- vector()

    #Run this loop as many times as specified by perm.cor
    for(i in 1:perm.cor){
      #Duplicate the d.sub matrix, and then replace entries with randomized numbers
      d.sub2 <- d.sub
      for(j in 1:dim(d.sub)[2]){
        #Set new seed for upcoming randomization
        set.seed(seeds[j + (i-1)*(dim(d.sub)[1]) + (m-1)*dim(d.sub)[2] ])

        #randomize each taxon's abundance vector
        d.sub2[, j ] <- sample(d.sub[ ,j ])
      }
      #replace focal column with original column
      d.sub2[, which.otu] <- d.sub[ , which.otu]

      #calculate correlations between randomized taxon vectors
      cor.mat.true.null <- cor(d.sub2)

      #save the vector corresponding to the focal taxon, m
      perm.cor.vec.mat <- cbind(perm.cor.vec.mat, cor.mat.true.null[,m])

    }
    #Take the median of the correlations produced by the null model
    med.otu.cors <- cbind(med.otu.cors, apply(perm.cor.vec.mat, 1, median))

    if(m %% 20 == 0){print(m)}
  }
}

#get observed - expected individual correlations
obs.exp.cors.each <- cor.mat.true - med.otu.cors
diag(obs.exp.cors.each) <- 0
```

```
#calculate connectedness values
conn.pos <- apply(obs.exp.cors.each, 2, pos.mean)
conn.neg <- apply(obs.exp.cors.each, 2, neg.mean)

#calculate cohesion values
#coh.pos <- rel.c %*% conn.pos
#coh.neg <- rel.c %*% conn.neg



###############################################################################
        ##################
###############################################################################
        ##################

##Start of analysis for analysing Bray-Curtis dissimilarity as a function of cohesion, calculated
        from different subsets of taxa

#Save original vectors of connectedness values
conn.pos.all <- conn.pos
conn.neg.all <- conn.neg

#create vectors to hold R2 values, p values, and parameter estimates for the regression model
coh.both.r2 <- vector()
pos.coh.p <- vector()
neg.coh.p <- vector()
pos.coh.sign <- vector()
neg.coh.sign <- vector()



##Calculate Bray-Curtis dissimilarity
#Use full abundance matrix (orig.rel)
bc <- as.matrix(vegdist(orig.rel))

#Take only samples within the specified boundaries of days between samples
days.elapsed.mat <- as.matrix(dist(as.numeric(as.Date(rownames(orig.rel), format =
        "%m/%d/%Y"))))
days.elapsed.mat[lower.tri(days.elapsed.mat)] <- 0
bc.diss <- bc[which(days.elapsed.mat >= elapse.min & days.elapsed.mat <= elapse.max, arr.ind
        = T)]
#find the samples that started off the gap
first <- rel.c[which(days.elapsed.mat >= elapse.min & days.elapsed.mat <= elapse.max, arr.ind =
        T)[,1], ]

par(mfrow = c(1,1))
```

```
#Make a vector with different numbers of taxa to keep
 how.many.keep <- ceiling(seq(dim(d.sub)[2], 4, by = -2))

 #Start loop to conduct the regression while keeping different numbers of taxa
 for(f in 1:length(how.many.keep)){

   #Repeat this portion for both positive and negative connectedness
   for(j in c( "positive", "negative")){
     #define the number of taxa to keep
     num.keep <- how.many.keep[f]
     #define whether working with positive connectedness or negative connectedness
     ifelse(j == "positive", conn.all <- conn.pos.all, conn.all <- conn.neg.all)
     #choose taxa to keep based on largest (by magnitude) connectedness values
    which.keep <- which(rank(abs(conn.all)) %in% seq(length(conn.all), length(conn.all) -
        num.keep + 1, -1) )

     #Create a new matrix "h" that contains only most connected taxa
     h <- first[, match( names(apply(d.sub, 2, mean)[which.keep]), colnames(first))]
     #Pull out connectedness values from only the most connected taxa
     conn <- conn.all[which.keep]
     #generate positive or negative cohesion vector
     ifelse(j == "positive", cohesion.pos <- as.matrix(h) %*% conn, cohesion.neg <- as.matrix(h)
        %*% conn)
   }

     #Save model R2 value, p values, and coefficient estimates
     coh.both.r2[f] <- summary(lm(bc.diss ~ cohesion.pos + cohesion.neg ))$adj.r.squared
     pos.coh.p[f] <- coef(summary(lm(bc.diss ~ cohesion.pos + cohesion.neg )  ))[2, 4]
     neg.coh.p[f] <- coef(summary(lm(bc.diss ~ cohesion.pos + cohesion.neg )  ))[3, 4]
     pos.coh.sign[f] <- coef(summary(lm(bc.diss ~ cohesion.pos + cohesion.neg )  ))[2, 1]
     neg.coh.sign[f] <- coef(summary(lm(bc.diss ~ cohesion.pos + cohesion.neg )  ))[3, 1]

 }

#Plot the adjusted R2 value of the model against the number of taxa included
par(mar = c(5, 5, 2, 2))
plot(coh.both.r2 ~ how.many.keep, ylim = c(0, .50),  ylab = expression("Model R" ^ 2), xlab =
        "Number of Taxa Included", type = "l", lwd = 3, lty = 1, cex.axis = 1.2, cex.lab = 1.5)

#Find the maximum adjusted R2 value
max(coh.both.r2)
#Find the number of taxa included at the max R2 value
how.many.keep[coh.both.r2 == max(coh.both.r2)]


#############################################################################
       #################
```

```
################################################################################
       #################

##Make a loop to run iterations of the above analysis with random taxa selected

#Make a vector to hold the R2 values of the model using different numbers of random taxa
coh.both.loop.r2 <- vector()

#determine number of iterations to run
iter <- 500

#create a matrix to hold the R2 values associated with each number of OTUs keps in each
       iteration
iter.r2.mat <- matrix(numeric(0), iter, length(how.many.keep))

for(runs in seq(1, iter, 1)){

  for(f in 1:length(how.many.keep)){

    #create a shuffled vector corresponding to the order of OTUs in the matrix
    rand.order <- sample(seq(1, length(conn.pos.all), 1))

    #decide how many OTUs to keep for the analysis
    for(j in c( "positive", "negative")){
      num.keep.rand <- how.many.keep[f]

      ifelse(j == "positive", conn.all <- conn.pos.all, conn.all <- conn.neg.all)

      #Select which taxa to keep in the analysis from the randomized vector
      which.keep.rand <- rand.order[1:num.keep.rand]

      #repeat subsetting of matrix and connectedness values using random taxa
      h.rand <- first[, match( names(apply(d.sub, 2, mean)[which.keep.rand]), colnames(first))]
      conn.rand <- conn.all[which.keep.rand]
      #generate cohesion vectors
      ifelse(j == "positive", cohesion.pos.rand <- as.matrix(h.rand) %*% conn.rand,
          cohesion.neg.rand <- as.matrix(h.rand) %*% conn.rand)
    }

    #save R2 of analysis using positive and negative cohesion to model BC dissimilarity
    #the vector represents one run using each different number of taxa (given in how.many.keep)
    coh.both.loop.r2[f] <- summary(lm(bc.diss ~ cohesion.pos.rand + cohesion.neg.rand )
        )$adj.r.squared

  }
```

```
  #save the vector of R2 values in a matrix
  iter.r2.mat[runs, ] <- coh.both.loop.r2

  if(runs %% 10 == 0){print(runs)}
}


#Run these lines to add in null model lines for mean and 5th / 95th quantiles
quant05 <- prop.vec(iter.r2.mat, .05)
quant95 <- prop.vec(iter.r2.mat, .95)

points(quant05 ~ how.many.keep, type = "l", lty = 3, lwd = 3, col = "red")
points(apply(iter.r2.mat, 2, median) ~ how.many.keep, type = "l", lty = 1, lwd = 3, col = "red")
points(quant95 ~ how.many.keep, type = "l", lty = 3, lwd = 3, col = "red")
```

**Appendix 6: Results of Human Microbiome Project Analyses (Column Shuffle)**



Anterior_nares



Regression Coefficients for Cohesion

# Attached_Keratinized_gingiva



# Regression Coefficients for Cohesion

**Buccal_mucosa**



**Regression Coefficients for Cohesion**

**Hard_palate**



**Regression Coefficients for Cohesion**

Left_Antecubital_fossa



Regression Coefficients for Cohesion

## Left_Retroauricular_crease



## Regression Coefficients for Cohesion

**Mid_vagina**

Adjusted Model $R^2$

Number of OTUs Included

**Regression Coefficients for Cohesion**

Slope Estimate

— Positive
— Negative

Number of OTUs Included

**Palatine_Tonsils**

**Regression Coefficients for Cohesion**

**Posterior_fornix**

*Adjusted Model R$^2$* vs *Number of OTUs Included*

**Regression Coefficients for Cohesion**

*Slope Estimate* vs *Number of OTUs Included*

Positive

Negative

**Right_Antecubital_fossa**

Adjusted Model $R^2$

Number of OTUs Included

**Regression Coefficients for Cohesion**

Slope Estimate

— Positive
— Negative

Number of OTUs Included

## Right_Retroauricular_crease



## Regression Coefficients for Cohesion

**Saliva**



**Regression Coefficients for Cohesion**

**Stool**

Adjusted Model R²

Number of OTUs Included

**Regression Coefficients for Cohesion**

Slope Estimate

— Positive
— Negative

Number of OTUs Included

## Subgingival_plaque



## Regression Coefficients for Cohesion

**Supragingival_plaque**

Adjusted Model R$^2$

Number of OTUs Included

**Regression Coefficients for Cohesion**

Slope Estimate

Positive
Negative

Number of OTUs Included

# Throat



Number of OTUs Included

# Regression Coefficients for Cohesion



Number of OTUs Included

**Tongue_dorsum**

Adjusted Model $R^2$

Number of OTUs Included

**Regression Coefficients for Cohesion**

Slope Estimate

Positive
Negative

Number of OTUs Included

**Vaginal_introitus**

**Regression Coefficients for Cohesion**

**Chapter 6: Perspectives and Future Work**

**Preface**

This thesis had two related objectives: apply ecological theory to varied empirical ecosystems, and use the insights learned from these studies to create more accurate models. This integration of ecological theory and microbial datasets appears to be particularly fruitful, because the theory provides a way to navigate the analysis of large, complex datasets. This chapter highlights intersections between microbial ecology and theoretical ecology, and it suggests some experiments and analyses that could build off of this dissertation.

**Introduction**

A major goal of microbial ecology is to predict microbial community dynamics, including rates of community turnover and eventual community composition (Fierer and Ladau 2012, Larsen et al. 2012). The hurdles to achieving this goal are both conceptual, such as identifying the drivers of community composition (Dini-Andreote et al. 2015), and statistical, such as developing methods to identify signals in high-dimensional, relative abundance data (Friedman and Alm 2012). Theoretical ecology has, as a field, dedicated substantial effort to quantifying and predicting community turnover, but little of this theory or statistical methodology has been used to model microbes. However, ecological theory has recently been applied in microbial systems as varied as the human microbiome (Morris et al. 2013) and the ocean (Fuhrman et al. 2015) to aid in the understanding of microbial community dynamics. Furthermore, several prominent ideas in theoretical ecology and microbial ecology address similar topics; for example, the microbial concept of identifying deterministic and stochastic

community processes has many parallels to the theoretical framework for analyzing the strength of stability in communities.

Partitioning the relative importance of deterministic vs. stochastic forces in microbial communities has been a topic of particular interest in recent years (Zhang et al. 2016). This idea is important to forecasting microbial community dynamics because it provides a framework for understanding when communities should be predictable and when communities are predisposed to erratic dynamics (Dini-Andreote et al. 2015). Some essential components of this research include quantifying inherent population variability in microbial systems, identifying factors that increase or decrease population variability, and determining the relative contribution of various drivers of variability (Bissett et al. 2013). These drivers of variability include both biotic (e.g. resource competition or predation) and abiotic (e.g. environmental disturbance or nutrient availability) forces (Stegen et al. 2015). However, these drivers of variability are often operating simultaneously, meaning that it is difficult to partition the influence of these forces.

The projects outlined here attempt to better partition the sources of community and population variability by 1) proposing analyses of existing samples to investigate the role that phylogenetic relatedness plays in regulating community turnover and 2) suggesting experiments that would separate the influence of biotic and abiotic forces.

**Proposed Future Projects**

*Project 1: Identifying keystone OTUs that indicate impending community shifts*

One practical application of Chapters 4 and 5 of this dissertation could be to help researchers identify key taxa that might indicate that microbial communities are going to either remain fairly compositionally stable through time *or* are prone to rapid compositional changes. For example, part of a water quality monitoring program might be to conduct weekly tests for the

abundance of 5 key OTUs that indicate the rate of compositional change in the microbial community. However, it would be optimal to test as few OTUs as possible. Thus, Project 1 would focus on ways to include information about phylogenetic relatedness to improve the workflow developed in Chapter 4. It would be expected that highly related taxa might also have respond similarly to environmental cues, which would cause correlations between taxa that do not reflect taxon interactions. Weeding out these environmentally-driven correlations would improve the signal of biotic interactions in the dataset, which could improve the results of the workflow. I formed this hypothesis from the observation that turnover in the Lake Mendota bacterial time series (temperate location) was not captured as well as in the San Pedro Ocean Time Series (subtropical location). Thus, the higher degree of seasonality in the Lake Mendota time series may contribute to spurious correlations between OTUs with similar niches due to stronger environmental forcing.

In the summer of 2015, I conducted fieldwork at Trout Lake to experimentally test the hypothesis (based on results from Chapter 5) that keystone taxa could predict major compositional changes before they occurred. I transplanted biofilm communities between the epilimnion and hypolimnion of Trout Lake and took samples every 2 days. There were treatments where communities permanently remained in the epilimnion (Epi) or hypolimnion (Hypo) of Trout Lake, as well as treatments where the communities were temporarily (Pulse) or permanently (Press) relocated from the hypolimnion to the epilimnion (Fig 1). I took a highly replicated time series (at least 9 samples per treatment at 9 time points) before, during, and after the experimental manipulations to be able to build microbial community networks of each treatment at each time point. Because of the strong differences in temperature, light, and dissolved oxygen between the epilimnion and hypolimnion, I would expect that the microbial

communities in the transplanted treatments changed substantially in community composition as a result of moving locations in the water column.

| | Time Point | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | T1 | T2 | T3 | T4 | T5 | T6 | T7 | T8 | T9 |
| **Epi** | E | E | E | E | E | E | E | E | E |
| **Pulse** | E | E | E | H | H | H | E | E | E |
| **Press** | E | E | E | H | H | H | H | H | H |
| **Hypo** | H | H | H | H | H | H | H | H | H |

*Figure 1:* Experimental design of the transplant disturbance experiment conducted in summer 2015. "E" represents communities located in the epilimnion of Trout Lake, and "H" represents communities located in the hypolimnion. Samples were collected every 48 hours.

Due to time limitations, these samples have not been analyzed. However, it would be very interesting to analyze them with 16S tag sequencing because of the phylogenetic information that would accompany these data. A first step would be to identify the OTUs that are most highly connected in the bacterial communities prior to the experimental disturbance (using the workflow in Chapter 4). Then, it would be interesting to test whether highly-related taxa show similar patterns of connectedness, with the expectation that closely-related OTUs would show more similar connectedness values than phylogenetically distant OTUs. Relatedness could be calculated using several different metrics of phylogenetic distance to assess which metric is best for the analyses. These values of relatedness could be incorporated into the existing workflow to produce an improved algorithm for identifying keystone taxa.

The experimental design employed in this project was specifically intended to answer questions about the resilience and recovery of communities. The post-disturbance time points could be used to determine whether 1) the abundances of these highly connected OTUs correspond to the rate of community composition change after the disturbance has been implemented and 2) whether these highly connected OTUs recover more or less quickly than the remainder of the OTUs when the communities are replaced in their original conditions. Under the hypothesis that the highly connected OTUs are important drivers of bacterial community turnover, I would expect these OTUs to exhibit changes in abundance before the rest of the bacterial OTUs begin to experience changes.

*Project 2: Partitioning biotic vs. abiotic forces in shaping community and population response to disturbance*

Prior work in my dissertation (Chapter 2) showed that environmental stressors decreased the OTU-level population variability within bacterial communities. One hypothesis for this pattern is that the environmental stress imposed a strong selective pressure on the taxa within the communities, thereby creating similar community composition by consistently favoring the same set of taxa. However, environmental stress can also alter

**A.** *H1: Selective pressure drives both phylogenetic and population variability*
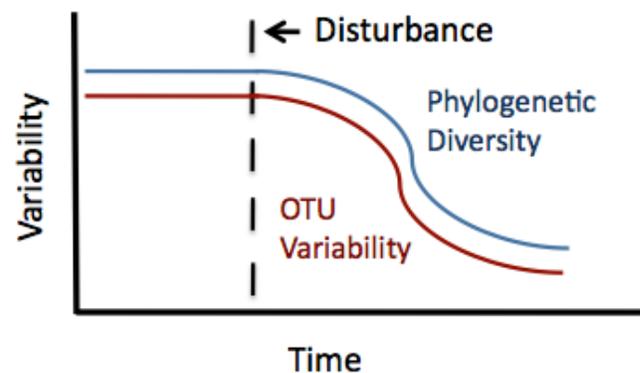
**B.** *H2: Post-disturbance biotic interactions drive population variability*
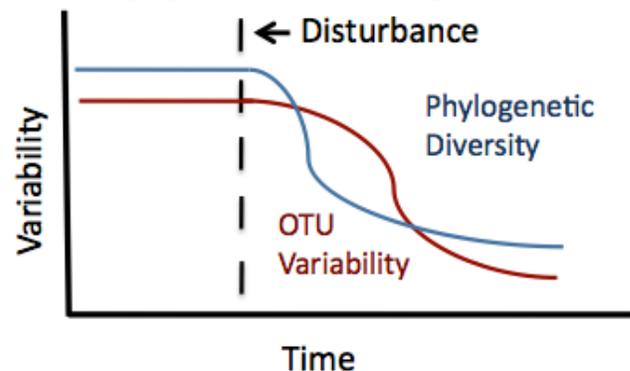
interactions between OTUs, meaning that biotic interactions could also contribute to the observed pattern. Currently, there are few methods to disentangle biotic vs. abiotic drivers of community composition in ecological communities. Again, the ability to obtain phylogenetic information from bacteria might make this question tractable in this system. Phylogenetic information would be an informative response variable in this case, because I would expect that closely related taxa would show similar responses to environmental selective forces, such that phylogenetic diversity would decrease following a disturbance (Helmus et al. 2010, Banks et al. 2013). Similarly, recent work has indicated that phylogenetic diversity changes in bacterial communities over successional trajectories (Brown and Jumpponen 2015). Thus, I would expect that, if selective pressure from abiotic stressors decreases OTU-level variability, phylogenetic diversity should show decreases at the same time points where OTU-level variability shows decreases (Fig. 2A). Conversely, if the drivers of OTU variability are the biotic interactions that occur after the phylogenetic restructuring, then OTU variability should decrease after phylogenetic diversity (Fig. 2B).

# References

Banks, S. C., G. J. Cary, A. L. Smith, I. D. Davies, D. A. Driscoll, A. M. Gill, D. B. Lindenmayer, and R. Peakall. 2013. How does ecological disturbance influence genetic diversity? Trends in Ecology & Evolution 28:670–679.

Bissett, A., M. V. Brown, S. D. Siciliano, P. H. Thrall, and M. Holyoak. 2013. Microbial community responses to anthropogenically induced environmental change: towards a systems approach. Ecology Letters 16:128–139.

Brown, S. P., and A. Jumpponen. 2015. Phylogenetic diversity analyses reveal disparity between fungal and bacterial communities during microbial primary succession. Soil Biology and Biochemistry 89:52–60.

Dini-Andreote, F., J. C. Stegen, J. D. van Elsas, and J. F. Salles. 2015. Disentangling mechanisms that mediate the balance between stochastic and deterministic processes in microbial succession. Proceedings of the National Academy of Sciences:201414261.

Fierer, N., and J. Ladau. 2012. Predicting microbial distributions in space and time. Nature Methods 9:549–551.

Friedman, J., and E. J. Alm. 2012. Inferring Correlation Networks from Genomic Survey Data. PLoS Computational Biology 8:1–11.

Fuhrman, J. A., J. A. Cram, and D. M. Needham. 2015. Marine microbial community dynamics and their ecological interpretation. Nature Reviews Microbiology 13:133–146.

Helmus, M. R., W. (Bill) Keller, M. J. Paterson, N. D. Yan, C. H. Cannon, and J. A. Rusak. 2010. Communities contain closely related species during ecosystem disturbance. Ecology Letters 13:162–174.

Larsen, P., Y. Hamada, and J. Gilbert. 2012. Modeling microbial communities: Current, developing, and future technologies for predicting microbial community interaction. Journal of Biotechnology 160:17–24.

Stegen, J. C., X. Lin, J. K. Fredrickson, and A. E. Konopka. 2015. Estimating and mapping ecological processes influencing microbial community assembly. Frontiers in Microbiology 6.

Zhang, X., E. R. Johnston, W. Liu, L. Li, and X. Han. 2016. Environmental changes affect the assembly of soil bacterial community primarily by mediating stochastic processes. Global Change Biology 22:198–207.