Cyanobacterial harmful algal blooms are a biological disturbance to western Lake Erie bacterial communities

Running title: Bacterial community ecology of CHABs

Michelle A. Berry[1], Timothy W. Davis[2], Rose M. Cory[3], Melissa B. Duhaime[1], Thomas H. Johengen[4], George W. Kling[1], John A. Marino[1], Paul A. Den Uyl[2], Duane Gossiaux[3], Gregory J. Dick[3,#], Vincent J. Denef[1,#]

[1]Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI, 48109. [2]NOAA Great Lakes Environmental Research Laboratory, Ann Arbor MI, 48108

[3]Department of Earth and Environmental Sciences, University of Michigan, Ann Arbor, MI, 48109. [4]Cooperative Institute for Limnology and Ecosystems Research, University of Michigan, Ann Arbor, MI, 48109.

[#]Corresponding authors: 1141 Kraus Natural Science, 830 N. University, Ann Arbor, MI 48109, vdenef@umich.edu, Phone: +1 (734) 764-6481, Fax: +1 (734) 763-0544; gdick@umich.edu, Phone: +1 (734) 763-3228, Fax: +1 (734) 763-4690

Author contributions: GJD, VJD, THJ, TWD, MBD, GWK, RMC designed the experiment, MAB, JAM, TWD, DG, PD performed experiments, MAB analyzed data, and MAB, VJD, GJD, GWK, RMC, TWD, JAM wrote the paper.

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version record. Please cite this article as doi:10.1111/1462-2920.13640.

Summary.

Human activities are causing a global proliferation of cyanobacterial harmful algal blooms (CHABs), yet we have limited understanding of how these events affect freshwater bacterial communities. Using weekly data from western Lake Erie in 2014, we investigated how the cyanobacterial community varied over space and time, and whether the bloom affected non-cyanobacterial (nc-bacterial) diversity and composition. Cyanobacterial community composition fluctuated dynamically during the bloom, but was dominated by *Microcystis* and *Synechococcus* OTUs. The bloom's progression revealed potential impacts to nc-bacterial diversity. Nc-bacterial evenness displayed linear, unimodal, or no response to algal pigment levels, depending on the taxonomic group. In addition, the bloom coincided with a large shift in nc-bacterial community composition. These shifts could be partitioned into components predicted by pH, chlorophyll *a*, temperature, and water mass movements. *Actinobacteria* OTUs showed particularly strong correlations to bloom dynamics. AcI-C OTUs became more abundant, while acI-A and acI-B OTUs declined during the bloom, providing evidence of niche partitioning at the sub-clade level. Thus, our observations in western Lake Erie support a link between CHABs and disturbances to bacterial community diversity and composition. Additionally, the short recovery of many taxa after the bloom indicates that bacterial communities may exhibit resilience to CHABs.

Originality-Significance Statement (not to appear in published manuscript).

CHABs are a global threat to freshwater resources. Case in point was western Lake Erie's 2014 CHAB that resulted in a drinking water shutdown in Toledo, Ohio, impacting 500,000 residents. Using weekly time-resolved molecular data, we describe the community ecology of *Cyanobacterial* taxa during this bloom, and we demonstrate how the bloom corresponded to shifts in bacterial diversity and composition. This work contributes to our understanding of how CHABs affect microbial communities, and it also contributes to a broader literature on disturbance and resilience of microbial communities.

Introduction.

Cyanobacterial harmful algal blooms (CHABs) are a major threat to freshwater ecosystems globally and are primarily driven by human activities (Paerl and Huisman, 2009; O'Neil *et al.*, 2012; Michalak *et al.*, 2013; Visser *et al.*, 2016). CHABs impact ecosystems and human health by diminishing habitat for plants and animals, disrupting food web dynamics, creating hypoxic zones, and producing toxins (Carmichael *et al.*, 2001; Conroy *et al.*, 2005; Hernández *et al.*, 2009; Miller *et al.*, 2010; Backer *et al.*, 2013). Despite a large body of CHAB research (Paerl and Otten, 2013; Steffen *et al.*, 2014; Davis and Gobler, 2016), relatively few studies have examined this phenomenon from a microbial ecology perspective that includes the community ecology of dominant cyanobacterial species as well as associations between cyanobacterial populations and other bacterial populations, which we will refer to as "nc-bacterial".

CHAB cyanobacterial diversity can vary both spatially and temporally within a lake. For example, a year-long study from Yanga Lake (Australia) found a succession of cyanobacterial consortia through time, which was determined by seasonal biotic and abiotic fluxes (Woodhouse *et al.*, 2015). In another example, a study from western Lake Erie (USA) found that *Microcystis* dominated in low P:N locations, while *Anabaena* and *Planktothrix* dominated in high P:N locations, because *Microcystis* was better able to scavenge phosphorus (Harke, Davis, *et al.*, 2016). While prior studies have investigated some of the spatiotemporal trends of CHAB communities, we lack insight into how these communities vary on highly resolved time scales. Increased temporal resolution of CHAB community datasets may elucidate additional ecological associations between CHAB species that are key to understanding bloom ecology.

Another important aspect of CHAB ecology is the extent to which these events impact nc-bacterial communities. We currently have poor understanding of if and how CHABs influence nc-bacterial richness and evenness (alpha diversity). Field studies from Lake Taihu (China) found no effect on bacterial alpha diversity (Tang *et al.*, 2010; Wilhelm *et al.*, 2011), while a study from Yanga Lake found that diversity increased with cyanobacterial bioviolume (Woodhouse *et al.*, 2015). These conflicting results could possibly be explained by differential responses between bacterial groups. In a study using pond mesocosms, the richness of bacterial groups were shown to have strikingly different responses to experimentally manipulated primary productivity measured by chl *a* (Horner-Devine *et al.*, 2003). Specifically, *Alphaproteobacteria* exhibited a negative unimodal relationship, *Bacteroidetes* exhibited a positive unimodal relationship, and *Betaproteobacteria* exhibited no relationship to chl *a* concentrations. Therefore, analysis

within taxonomic groups may help clarify the influence of CHABs on nc-bacterial alpha diversity.

Similarly, CHABs are known to influence the composition of bacterial communities, but again, prior studies have reported conflicting results. A study from Lake Taihu found that bacteria attached to organic aggregates were different between two sites with differing chl *a* concentrations (Tang *et al.*, 2010), but they also reported strong influences of co-varying factors such as temperature, oxygen, turbidity, and inorganic nutrients. Meanwhile, a study from Yanga Lake reported that bacterial community composition was influenced by pH, temperature, oxygen, and conductivity during a CHAB (Woodhouse *et al.*, 2015). Therefore, the relative impacts of CHABs versus abiotic factors on nc-bacterial community composition are still unclear.

To address these outstanding questions in CHAB microbial ecology, we investigated spatiotemporal dynamics of cyanobacterial populations, as well as changes to nc-bacterial alpha diversity and composition during the 2014 western Lake Erie CHAB. Lake Erie is the twelfth largest freshwater lake on Earth by surface area (Ohio Department of Natural Resources). It also provides essential ecosystem services by supporting a $1 billion USD fishing economy and supplying drinking water to over 11 million people (Ohio Department of Natural Resources; Bingham *et al.*, 2015). The CHAB in 2014 is of particular relevance, because it led to the drinking water crisis in Toledo, Ohio (Tanber, 2014). Insights into freshwater bacterial community ecology during CHAB events can point us towards possible interactions between cyanobacterial species that govern CHAB development and termination, and it can also inform

predictions of nc-bacterially-mediated ecosystem processes during these high impact disturbances.

Results and Discussion.

Bloom diversity, toxicity, and ecology

A CHAB occurred in western Lake Erie between late July and late October of 2014, characterized by elevated algal pigments (chl *a* and phycocyanin) and elevated particulate microcystin cyanotoxins (Figure 1b). We observed the bloom at three sites: the two nearshore sites, situated near the Maumee River, had higher median chl *a* concentrations than the offshore site (Figure 1a; Nearshore1 median: 18.5, Nearshore 2: 13.72, Offshore median: 5.86). However, the range of pigment values at each site was large, so on certain dates e.g., first August time point (Aug. 4) and first and second September time points (Sept. 2, Sept. 8), the offshore site had similarly high levels as nearshore sites. In later analyses, we leverage these temporal differences in bloom intensity between nearshore and offshore sites to model variation in nc-bacterial composition associated with the bloom.

Chl *a* and phycocyanin concentrations measured at the same site and date were highly correlated ($p < 0.001$, Spearman's rho = 0.793). Correlations between time-series can result in spurious results (Johansen, 2007), but visual inspection indicated that the two variables were qualitatively correlated (Figure 1b). Since chl *a* is produced by most phytoplankton, but phycocyanin is only produced by *Cyanobacteria*, this analysis suggests that *Cyanobacteria* dominated the bloom dynamics. However, these data do not preclude the presence of eukaryotic phytoplankton species. In fact, our universal 16S

primers picked up numerous chloroplast reads, suggesting that eukaryotic species were present. Eukaryotic algae were not the focus of this study, so they were removed from the dataset.

Particulate microcystin toxin was correlated with phycocyanin concentrations ($p < 0.001$, Spearman's rho = 0.836), but this relationship differed qualitatively between early and late bloom periods (Figure 1b). From mid-July to late August, elevated phycocyanin corresponded to high levels of particulate microcystin. Then from early September to October, elevated phycocyanin corresponded to lower toxin concentrations. Despite this shift in toxicity, there was a single dominant *Microcystis* OTU present in the community (Figure 2b). These data can be explained by the fact that there are numerous *Microcystis* strains that have more than 97% similarity in their full-length 16S rRNA gene, yet differ with respect to toxigenic potential and other ecological traits (Harke, Steffen, *et al.*, 2016).

The cyanobacterial community was a diverse community of 11 non-rare OTUs (mean relative abundance > 0.05 %) assigned to *Synechococcus*, *Microcystis*, unclassified genera, *Pseudanabaena*, and *Anabaena* (in order of mean relative abundance) (Figure 2a, 2b). Prior studies of CHABs on Lake Erie have primarily used microscopy to identify cyanobacterial species, and have reported that *Microcystis* was the heavily dominant cyanobacterium by biomass (Bridgeman *et al.*, 2013; Michalak *et al.*, 2013; Steffen *et al.*, 2014; Harke, Davis, *et al.*, 2016). In contrast, our molecular data indicates that the 2014 CHAB consisted of a more diverse cyanobacterial community that varied highly in composition over space and time. *Synechococcus* initiated the bloom at all three sites, and remained an abundant genus (by gene copy abundance) throughout the entirety of the

bloom. However, once *Microcystis* rose to abundance (3-4 weeks after *Synechococcus*), it dominated at the nearshore stations, while *Synechococcus* continued to dominate at the offshore station (Figure 2a).

Only a few studies have discussed the abundance and importance of *Synechococcus* in Lake Erie's CHABs (Ouellette *et al.*, 2006; Gobler *et al.*, 2008; Davis *et al.*, 2012). This discrepancy could be related to either a predefined focus on *Microcystis* (e.g., using *Microcystis* specific primers), or a bias against picobacteria during sampling and morphological identification (e.g., colony-only sampling or a focus on the higher biomass per cell of colonial and filamentous *Cyanobacteria*). However, *Synechococcus* species likely co-occur with *Microcystis* in several systems, because a molecular-based study on Lake Taihu found that *Synechococcus* was abundant during *Microcystis* blooms (Ye *et al.*, 2011). We observed both positive and negative correlations between *Microcystis* and *Synechococcus* OTUs in our study, though only one correlation between *Microcystis* and *Synechococcus* OTU 177 was significant (Table S1). Still, the persistent dominance of *Synechococcus* at the offshore station, where *Microcystis* abundance was generally lower, suggests that there may be a competitive or antagonistic interaction between these taxa. In fact, microcystin has been shown to inhibit the growth of some *Synechococcus* species (Hu *et al.*, 2004). Future experimental studies might address how ecological interactions with *Microcystis* vary among different *Synechococcus* taxa.

Pseudanabaena was the third most abundant genus, and like *Microcystis*, we only detected one abundant OTU (Figure 2b). *Pseudanabaena* co-occurred with *Microcystis* throughout much of the bloom, and the relative abundances of the two were highly

correlated (Table S1). Previous studies have shown that *Pseudanabaena* can be epiphytic on *Microcystis* colonies (Agha *et al.*, 2016), therefore it is not surprising that these two genera would be correlated. Furthermore, some *Pseudanabaena* strains have the genetic potential to produce cyanotoxins (Rangel *et al.*, 2014), thus co-occurrence of *Pseudanabaena* and *Microcystis* may have repercussions for CHAB toxicity.

Finally, observed dynamics of *Anabaena* and *Microcystis* support prior hypotheses addressing the competitive advantages of either genus under different nutrient regimes. *Anabaena* and *Microcystis* were not significantly anti-correlated (Table S1), but their opposing relationship was qualitatively apparent. For example, *Anabaena* was mostly absent from the Erie basin in the first stage of the bloom when phosphorus concentrations were low (Figure 2a, Figure S1). However, in late summer, when phosphorus levels increased and dissolved inorganic nitrogen concentrations decreased, *Anabaena* reached its peak level relative to *Microcystis*, particularly at the offshore site. These patterns are consistent with previous findings that *Microcystis* upregulates P-scavenging genes and outcompetes *Anabaena* in P-limited environments (Gobler *et al.*, 2016; Harke, Davis, *et al.*, 2016).

One major caveat to the observations described above, as well as correlations made between taxa in the rest of this paper, is that these associations are biased in several ways by extraction protocol, primers, and the compositional nature of sequence data (Aitchison, 1982; Brooks *et al.*, 2015). The universal primer set we used is known to be biased against SAR11, an *Alphaproteobacteria* common in marine environments, which has a sister lineage (LD12) that is ubiquitous in freshwater systems (Apprill *et al.*, 2015). Furthermore, a correlation does not imply causality or even interaction. Still, these

observations can be quite informative when considered in the context of other observational and laboratory data.

In summary, the cyanobacterial community in western Lake Erie's 2014 CHAB was diverse and highly dynamic at the OTU level. The community was spatially heterogeneous, such that *Synechococcus* and *Anabaena* were more abundant offshore and *Microcystis* dominated near to shore. In addition, weekly temporal sampling revealed putative associations between cyanobacterial taxa, which could form the basis for more specific experimental studies examining pairwise ecological interactions.

Seasonal and bloom-associated patterns in nc-bacterial alpha diversity

Nc-bacterial richness and evenness exhibited differing temporal dynamics during the bloom cycle. With the exception of a few highly variable samples in October, observed nc-bacterial richness increased throughout the season (Figure 3A). In contrast, nc-bacterial evenness, measured by Simpson's E, decreased until October (Figure 3E). It should be noted that rarefaction curves of OTU richness rarely reach saturation in diverse microbial environments, so richness estimates are highly dependent on sequencing depth. Therefore, we reported our estimates as observed richness (out of 15,631 sequences) rather than true richness. Still, our richness estimates show consistent trends with other studies that have observed increasing bacterial diversity from the spring to early fall in freshwater systems (Shade *et al.*, 2007; Kara *et al.*, 2012) and in the surface waters of marine systems (Cram *et al.*, 2015).

We did not observe a relationship between algal pigments and nc-bacterial richness (Figure 3B-D, Figure S2); however, we did find relationships between algal

pigments and the evenness of certain taxonomic groups (Figure 3F-H, Figure S2). *Alphaproteobacteria* evenness exhibited a unimodal response to log chl *a*, while *Bacteroidetes* evenness exhibited a linear response. The evenness of *Betaproteobacteria* was also slightly positively correlated with log chl *a*, though the association was not strong enough to be certain. However, the Inverse Simpson Index, which combines both richness and evenness, showed a much stronger response for linking log chl *a* to *Betaproteobacteria* ($p < 0.001$, $R^2 = 0.328$). Therefore, this analysis is quite sensitive to the measure of alpha diversity used.

In general, our data suggest that the bloom influences bacterial evenness more than bacterial richness. We hypothesize that increases in the evenness of dominant bacterial groups during the CHAB could be related to an increase in habitat complexity (colony-attached communities) or substrate complexity (carbon compounds from a diverse algal community), which would allow rare or dormant taxa to become relatively more abundant. While bloom specialists might be expected to dominate during this period, rapid weekly shifts in algal abundance (assumed from changes in pigments) and cyanobacterial composition could inhibit this, thereby promoting a more even community.

In addition to chl *a*, we investigated the relationship of nc-bacterial richness and evenness with other measurements of the bloom. Lake pH can increase to very high levels during cyanobacterial blooms due to heightened primary productivity (López-Archilla *et al.*, 2004), because photosynthesis fixes carbon and displaces the equilibrium of carbon dioxide/bicarbonate/carbonate that would otherwise buffer a freshwater system. Compared to chl *a*, pH showed very similar and slightly stronger trends with respect to

evenness of *Alphaproteobacteria* and *Betaproteobacteria* (Figure S3). Chl *a* is often used as a proxy for primary productivity (Downing and Leibold, 2002; Horner-Devine *et al.*, 2003; Smith, 2007), but light, nutrients, and grazing rates can decouple the two by affecting per cell concentrations of chlorophyll or by reducing the standing stock of phytoplankton (Behrenfeld *et al.*, 2005). Lake pH can be affected by geochemical conditions e.g., salt concentrations and presence of mineral carbonates, but there is no evidence for these conditions changing rapidly in western Lake Erie during the summer season. Therefore, pH might be a better proxy for primary productivity than chl *a* in this system, and would consequently correspond more strongly to bacterial diversity if such a relationship exists. Phycocyanin showed similar trends to nc-bacterial evenness (Figure S3), but the relationships were weaker, which suggests that nc-bacterial evenness is more affected by the total algal community than solely *Cyanobacteria*.

Our data, supporting a link between the bloom and evenness of certain bacterial taxa, are consistent with experimental evidence that primary productivity affects alpha diversity of bacterial groups in different ways (Horner-Devine *et al.*, 2003). However, the actual relationships we observed were quite distinct for each taxonomic group. Specifically, *Alphaproteobacteria* exhibited a U-shaped response to chl *a* in a pond mesocosm study (Horner-Devine *et al.*, 2003), but our study shows the inverse hump-shape. The mesocosm study found a hump-shaped response for *Betaproteobacteria*, but we found a positive linear trend. Discrepancies between these studies could be due to differences in community composition, differences in the range of chl *a* levels over which the communities were sampled, or other environmental factors that differ between a field and lab environment. Our results also differ from other CHABs field studies that have

found no effect of the bloom on bacterial alpha diversity (Eiler and Bertilsson, 2004; Woodhouse *et al.*, 2015), though these studies only examined total bacterial richness. In lieu of our findings, future studies should examine both bacterial richness and evenness, and should explore diversity patterns within major taxonomic groups.

This study provides some initial data differentiating between how annual cycles and bloom-associated trends affect the richness and evenness of freshwater nc-bacterial groups. Future studies that expand our observations across multiple years and cover the full annual range of seasonal variation will further resolve the intertwined effects of seasonality and CHABs growth dynamics on bacterial alpha diversity.

Influence of CHABs and abiotic seasonal factors on nc-bacterial community composition

The nc-bacterial community exhibited large shifts in composition during the 2014 bloom cycle. The Bray-Curtis dissimilarities between the first June samples and peak bloom dates in August or September were 0.784, 0.812, and 0.642 for nearshore1, nearshore2, and offshore respectively. We expected that several biotic and abiotic factors contributed to these fluctuations, so rather than examining several factors independently, we used principal coordinates to identify the major axes of variation within the community. We then determined which variables corresponded to change across each axis over time with linear time-series models. In considering each principal coordinate, we examined whether sample scores were similar between nearshore and offshore sites. Abiotic seasonal dynamics should influence all three stations similarly, while the CHAB, if it has an effect, should influence the offshore site differentially than the nearshore sites on dates with large discrepancies in bloom intensity. We considered three principal

coordinates, because the third coordinate was situated at an obvious inflection point of the scree plot for relative variance explained by each axis (Figure S4).

The first principal coordinate (PC1) of Bray-Curtis sample dissimilarities explained 34.8% of variation in nc-bacterial community composition across all samples. PC1 scores exhibited a hump-shaped response over time, which was highly consistent for the two nearshore sites, but showed differences between nearshore and offshore sites in mid August and mid to late September (Figure 4A). These differences corresponded to dates when algal pigments were considerably lower at the offshore station than nearshore stations (Figure 1B), suggesting that the bloom could be an influencing factor. We attempted to model PC1 scores solely with environmental data, but we achieved much better results when time was included as an additional covariate. The best model included time and pH, though the model with time and chl *a* had a similar R^2 value (Table 1). For the model including pH, residuals were normal and did not exhibit autocorrelation (Figure S5), indicating that model assumptions were met. Model cross-validation returned a low mean squared error, indicating that the model was highly predictive.

We posit that PC1 reflects changes in composition associated with the bloom and other seasonal factors. pH and chl *a* were the two strongest environmental predictors of PC1, and they can both serve as measures of bloom intensity. pH increases during blooms because algal photosynthesis removes carbon dioxide from the water and increases hydroxide ion concentration. In our sampling season, pH reached exceedingly high levels for a lake (>9, Figure S2), which is indicative of very high primary productivity in an otherwise well-buffered system (López-Archilla *et al.*, 2004). Our model also suggests that seasonal variation is important, because the model fit improved considerably when

time was added as a covariate. Due to the limited interval of our study, it is difficult to interpret the meaning of the time variable. We think it's likely that there is a sinusoidal seasonal trend, but it appears as a linear trend during this four-month period.

There are multiple mechanisms by which an algal bloom can affect bacterial community composition. A shift from allochthonous to autochthonous dissolved organic carbon was observed during this CHAB (Cory *et al.*, 2016), which may have influenced the relative abundance of different taxa. Several other studies have documented that bacterial communities respond to shifts in substrates available within the dissolved organic carbon pool during both freshwater and marine blooms (Lau *et al.*, 2007; Teeling *et al.*, 2012; Yang *et al.*, 2015). Alternatively, pH is known to be a major influence on bacterial community composition in soil (Lauber *et al.*, 2009) and freshwater systems (Lindstrom *et al.*, 2005; Llirós *et al.*, 2014). Therefore, the bloom may have actually influenced the composition of nc-bacterial communities by changing the lake's pH. pH was found to be the most important factor structuring bacterial communities across 15 North European lakes spanning the range of 5.5 to 8.7 (Lindstrom *et al.*, 2005), and in Tibetan lake sediments spanning the range of 6.88 to 10.37 (Xiong *et al.*, 2012). The pH range in our study spanned from 7.9 to 9.3, which is smaller than other studies, but may have covered critical thresholds. Importantly, the correlation of pH with PC1 suggests that changes to community composition were not driven solely by the presence of *Cyanobacteria* or harmful cyanobacterial species, but rather by the cumulative properties of the bloom, which would include eukaryotic or non-harmful species. If this is true, CHABs may not be particularly distinct disturbances to bacterial communities from other phytoplankton blooms that reach the same magnitude of primary productivity.

The second principal coordinate represented 11.0 % of the variation in nc-bacterial community composition, which was less than one-third of the variation explained by PC1. Unlike PC1, sample scores on PC2 were highly similar between nearshore and offshore sites on all dates (Figure 4C). Therefore, it is unlikely that bloom-related factors were strongly correlated to this axis of variation. Temperature was the best predictor of PC2 scores (Table 1). We did not include time as a covariate, because time was highly correlated with temperature, and it created multicollinearity issues in our model. The temperature model residuals were normal, and did not exhibit significant autocorrelation (Figure S5). However, cross-validation returned a somewhat high mean squared error, indicating that model estimates could still be biased by temporal trends for which we didn't account. Nevertheless, our model supports that temperature was an important factor in the structuring of nc-bacterial community composition. Congruently, freshwater bacterial communities are known to undergo seasonal shifts, and temperature has been found to be the single largest determinant of these patterns (Kent *et al.*, 2004; Crump and Hobbie, 2005; Shade *et al.*, 2007).

Finally, PC3 explained only 6.72% of variation in nc-bacterial community composition. PC3 scores differed strongly between nearshore and offshore sites, but unlike PC1, these differences did not correspond to dates with large discrepancies in bloom intensity (Figure 4E). Conductivity was the best predictor of PC3 scores (Table 1), and most model assumptions regarding normal independent residuals were met, though there was some autocorrelation in the residuals from nearshore1 (Figure S5). The Detroit and Maumee rivers are known to have distinct conductivity signatures (Millie *et al.*). Therefore, we interpret variation on this third axis as driven primarily by differences in

water mass, which result from differential inputs of the Maumee and Detroit rivers to nearshore and offshore sites.

Thus, using three principal coordinates, which together explain more than half of the variation in nc-bacterial community composition, we identified pH, chl *a*, temperature, and conductivity as key environmental gradients. PC1, which explained more than three times the variance of the second and third coordinates, showed evident differentiation between nearshore and offshore sites on dates with large differences in pH and chl *a*. Therefore, we argue that the bloom was a considerable disturbance to nc-bacterial community composition.

Bloom effects on abundant nc-bacterial groups and resilience of bacterial communities to CHAB disturbances

Principal coordinates analysis revealed that bloom-associated measures corresponded to changes in nc-bacterial community composition, but it did not reveal which taxa were most affected. Therefore, we investigated which nc-bacterial taxa significantly correlated with shifts in pH and chl *a*. Using Spearman's rank correlation tests, we found 34 abundant OTUs (mean relative abundance > 0.1%) that were positively correlated with pH and 27 that were negatively correlated (Table S2). There was considerable overlap (83%) with the OTUs associated with chl *a*. A majority of the most significant positive and negative correlated taxa to both bloom measures were *Actinobacteria* acI OTUs. *Actinobacteria* acI was also the most abundant clade in the nc-bacterial community by at least three-fold. Interestingly, acI-A and acI-B OTUs decreased during the CHAB, while acI-C OTUs increased (Figure 5). In addition,

changes in the relative abundance of these OTUs differed between nearshore and offshore stations, particularly on dates in mid August when there was a large discrepancy in algal pigments and pH. These data suggest there was niche partitioning among acI OTUs in response to the CHAB, which was conserved at the sub-clade level. Numerous other studies have documented niche and seasonal partitioning patterns in acI sub-clades (Allgaier and Grossart, 2006; Newton *et al.*, 2011; Eiler *et al.*, 2012), including partitioning by the ratio of allochthonous to autochthonous carbon (Jones *et al.*, 2009) as well as by pH (Newton *et al.*, 2007). However, these prior studies focused on partitioning between acI-A, acI-B, acII, and acIV. We found no published research on the ecology of acI-C, so further work will be necessary to determine the mechanism by which this sub-clade benefits from CHABs, and whether this mechanism is distinct from non-CHAB algal blooms.

Other abundant clades such as bacI, betI, bacV, and betIV did not show the same conserved niche partitioning to the bloom as acI (Figure S7). Within each clade, there were individual OTUs that appeared to respond positively or negatively during the bloom, but there were also abundant OTUs whose relative abundance did not strongly reflect bloom dynamics.

Dynamics at the OTU level, particularly among the acI, demonstrate that nc-bacterial community composition was highly affected by the western Lake Erie CHAB. Thus, bacterial communities exhibit a high degree of sensitivity to CHAB disturbances (Shade *et al.*, 2012). Nevertheless, by the end of October, acI OTUs recovered towards pre-bloom relative abundance. Similarly, PC1 scores returned nearly to pre-bloom levels and Bray-Curtis dissimilarities between the first and last time points at each site were

substantially smaller than the peak levels observed during the bloom (nearshore1: 0.460, nearshore2: 0.453, offshore: 0.364). This quick recovery toward baseline levels indicates community resilience (Shade et al. 2012). Freshwater bacterial communities have previously been shown to be highly resilient to physical and chemical disturbances (Shade *et al.*, 2011), and our study indicates that the same may be true for biotic disturbances.

Conclusion.

Western Lake Erie's bacterial community exhibited changes in diversity and composition during the bloom season of 2014. In particular, the evenness of *Alphaproteobacteria* and *Betaproteobacteria* showed differential responses to algal pigment levels, suggesting that the bloom affected niche diversity for these phylogenetic groups. Changes in community composition could be represented in three coordinates, with the first coordinate associated most strongly with bloom measures, the second coordinate associated with temperature, and the third coordinate associated with physical water mass movements. These results support work by others demonstrating that bacterial communities are impacted by CHABs, and identifies the acI clade as a particularly affected group. The time resolution of this study also demonstrates that most taxa affected by the CHAB exhibit resilience by recovering to pre-bloom levels shortly after the termination of this biological disturbance. A better understanding of the specific relationships and processes between bacterial diversity and the occurrence and toxicity of CHABs will be useful given the projected acceleration of CHABs in future years.

Experimental Procedures.

Sample collection

Samples were collected approximately weekly between mid-June and late October, 2014 from three stations (nearshore1, nearshore2, offshore) in the western basin of Lake Erie that correspond to NOAA Great Lakes Environmental Research Laboratory long-term monitoring sites WE12, WE2, WE4 respectively (NOAA-GLERL). Nearshore1 is closest to the water intake for the city of Toledo, nearshore2 is close to the mouth of the Maumee River, and the offshore site is on the northeastern edge of the bloom perimeter (Figure 1a).

Physicochemical measurements and microbial samples were obtained from an integrated 20 L water sample taken between the surface and 1 m above the bottom. The sample was homogenized by shaking. All station depths ranged between 4-12 m, and the shallowness of the western basin prevents vertical stratification. Temperature, pH, and conductivity were measured on deck, and algal pigment, cyanotoxin, and nutrient measurements were analyzed at NOAA-GLERL using standard techniques (U.S. EPA, 1979). H_2O_2 measurements were analyzed according to Cory et al. (2016). For microbial samples, a 2 L subsample was taken from the 20 L sample and rehomogenized. 150 mL was syringe filtered onto a 0.22 μm Millipore Express Plus filter (EMD Millipore, Billerica, MA), though on peak bloom dates the filter clogged before the full volume was filtered. All filter samples were placed into cryovials with 1 ml of RNAlater (Ambion, Foster City, CA) and frozen at -80 °C until extraction.

DNA Extraction and Sequencing

Filters were thawed at room temperature and, while folded with biomass facing inwards, rinsed with sterile PBS to remove RNAlater preservative. Filters were incubated in 100 µL Qiagen ATL tissue lysis buffer, 300 µL Qiagen AL lysis buffer, and 30 µL proteinase K for 1 hour at 56 °C on a rotisserie (Qiagen, Hilden, Germany). Cells were lysed by vortexing for 10 minutes. Lysates were homogenized with the Qiashredder column, and DNA was purified from the filtrate using the Qiagen DNeasy Blood and Tissue kit according to standard protocol. Extracted DNA was amplified using primer set 515f/806r, which targets the V4 hypervariable regions of the 16S rRNA gene (Bergmann *et al.*, 2011). The DNA was then sequenced using Illumina MiSeq v2 chemistry 2x250 (500 cycles) at the University of Michigan Medical School. RTA v1.17.28 and MCS v2.2.0 software were used to generate data. Fastq files were submitted to the NCBI sequence read archive under BioProject PRJNA318386, SRA accession number SRP07334.

Sequence Filtering and Pre-processing

Mothur V 1.34.3 was used to perform quality control and cluster sequences into OTUs (Schloss *et al.*, 2009). Sequence processing was performed according to the Mothur standard operating procedure (http://www.mothur.org/wiki/MiSeq_SOP accessed on March 13, 2015). Taxonomy was assigned to sequences using the Wang method (Wang *et al.*, 2007) with an 80% bootstrap cutoff using the Freshwater Microbial Field Guide (FWMFG) (Newton *et al.*, 2011). This database resolves clade and sub-clade level taxonomy for common freshwater taxa and allows our data to be compared to other freshwater studies. However, the FWMFG lacks certain taxonomic groups such as

Planctomycetes, so we used the Silva database V119 (Quast *et al.*, 2013) for the remaining unassigned reads. OTUs were clustered using the average neighbor algorithm with a 97% similarity threshold. Mothur output files were imported into R V 3.2.2 (R Core Team, 2015) using the phyloseq package V 1.10 (McMurdie and Holmes, 2013) for all downstream analyses of diversity and community composition. All scripts, mother output files, and sample data are publically available at https://github.com/DenefLab/chab-ecology.

Spatial and temporal bloom dynamics

Spearman's correlation tests were used to determine if there were monotonic relationships between algal pigments, toxin, and pH. To explore positive and negative associations between Cyanobacteria OTUs, we performed pairwise Spearman's correlation tests between all OTUs with mean relative abundance > 0.0005 using the corr.test command in the psych package with fdr correction (Revelle, 2015).

Bacterial alpha diversity

Alpha diversity was estimated using observed OTU richness and Simpson's Evenness (Simpson's E), which is the Inverse Simpson's Index divided by richness. Alpha diversity estimates were calculated for each sample by sampling sequences with replacement to 15,631 reads (the smallest library size) and averaging the measures over 100 trials using the estimate_richness command in phyloseq. Based on scatterplot visualization, we ran either linear or polynomial models to predict the richness and evenness of different bacterial groups from log chl *a* concentrations. Chl *a* measurements

were log scaled in order to meet assumptions of normal residual terms. P-values for each model were adjusted with a Benjamini-Hochberg false discovery rate (FDR) correction.

Bacterial community composition analyses

Differences in nc-bacterial community composition were calculated using the Bray-Curtis dissimilarity. Before calculating Bray-Curtis, data was transformed by scaling the raw proportions of OTUs to the read count of the smallest library (15,631 reads in this study), and rounding to the nearest integer count. This method is equivalent to the estimated value of averaging counts from repeated rarifying trials, but is more reproducible and does not contribute additional noise to the dataset (McMurdie and Holmes, 2014). The relative abundance of an nc-bacterial OTU was measured with respect to the nc-bacteria rather than the total bacterial community to reduce bias from changes in the cyanobacterial community. However, this method does not completely eliminate compositional effects (Aitchison, 1982).

To investigate differences in nc-bacterial community composition, we implemented a principal coordinates analysis. The goal of this analysis was to visualize similarity between samples in reduced dimensions, and to identify the major axes of variation in community composition through time. These axes are likely, though not certain, to correspond with environmental gradients. PCoA and related eigen-analyses have been implemented with time series data (Freeman *et al.*, 2014; Maurice *et al.*, 2015), and the interpretation is similar to other datasets except the sample scores on each axis are ordered by time. The percentage of variance explained by each axis was determined from the axis eigenvalue divided by the cumulative sum of all eigenvalues.

As the Bray-Curtis dissimilarity is non-euclidean, some principal coordinates (PCs) had negative eigenvalues, so we applied a Lingoes correction (Lingoes, 1971). To determine the number of principal coordinates to examine, we looked for the inflection point in the scree plot, which displays the relative variance in Bray-Curtis dissimilarity explained by each coordinate.

To investigate gradients that could be associated with changes in bacterial community composition, we constructed linear models using environmental variables (e.g. nutrients, temperature, pigments) to predict Bray-Curtis principal coordinate scores. Time series often contain long-term trends in addition to short-term fluctuations. Therefore, we experimented with including time as an additional covariate in each model. We assumed that differences between nearshore and offshore sites were due to environmental conditions, rather than inherent differences between these sites, so our models only included fixed effects. Model residuals were examined to determine whether they met assumptions of normality and independence (i.e. no autocorrelation). To assess model accuracy, we performed "leave-one-timepoint-out" cross validation of the best models for each axis and reported the mean squared error. This protocol is similar to LOOCV, except rather than removing one sample at a time during the model training stage, we removed all three samples from a given time point. This provided a less biased estimate of model error, because measurements from the same dates were frequently similar across sites, and would have reduced the error on the test set.

In addition to the simple linear models, we attempted to model each set of principal coordinates scores with multiple linear regression. We included all environmental variables as potential covariates and used best subset selection to identify

the model that minimized the Bayesian Information Criterion. The BIC penalizes more complex models in order to optimize the total amount of variance explained while reducing variables that contribute little explanatory power. Because we had relatively few data points (53) and many potential covariates (13), we also implemented a bootstrap analysis, in which we sampled with replacement and refit the models 100 times in order to determine the stability of a particular predictor. However, even with the bootstrapping, the results of each model varied depending on the seed value. We found that there was really only one stable predictor (present in >90% of all bootstrapped models), which is why we proceeded with simple time series models that included a single environmental covariate.

The principal coordinates approach that we took is an example of an indirect gradient analysis – gradients are unknown a priori and are estimated by linking environmental variables to the canonical axes of a sample similarity measure. We also tried an implementation of direct gradient analysis using redundancy analysis (RDA). RDA is a constrained version of principal components analysis in which the canonical axes are linear combinations of the response variables and also relate to the response variable via multiple linear regression. A time series version of RDA can be implemented using asymmetric eigenvector maps (AEM) (Baho *et al.*, 2015). Our RDA results identified the same gradients as our PCoA approach among others. Ultimately, we found the PCoA approach to be more appropriate, because the variance explained by the constrained axes was not much more than the unconstrained axes, indicating that the model was missing some important environmental gradients. In particular, we found it

more accurate and intuitive to observe the behavior of nearshore vs. offshore sites over time in unconstrained ordination space than constrained ordination space.

Finally, we performed the same principal coordinate analysis and time series linear model approach for the cyanobacterial community. However, the principal coordinates exhibited very noisy trends over time and were not strongly correlated to specific environmental variables. We also had concerns that the compositional nature of the OTU counts would lead to stronger biases in these analyses, because the cyanobacterial community constituted a relatively small proportion of all bacterial reads. Therefore, we report more descriptive statistics of the cyanobacterial community over time rather than implementing a model-based approach.

OTU-level analysis

To find potential positive or negative associations between nc-bacteria and the bloom, we performed Spearman's rank correlation tests with pH and chl *a*. We examined all nc-bacteria OTUs with mean relative abundance larger than 0.1% (107 taxa total) using the corr.test command from the psych package (Revelle, 2015) with an FDR correction.

Acknowledgments.

This work was supported by a grant from the Erb Family Foundation made through the University of Michigan Water Center and by the Great Lakes Restoration Initiative. RMC was supported by NSF CAREER award 1351745. We are grateful to the crew from NOAA-GLERL who assisted with fieldwork. We thank current members of the Denef,

Dick, and Duhaime labs for discussion of the manuscript. We also thank Prof. Kerby Shedden (University of Michigan) for in depth discussion and consultation on statistical methods. This manuscript is NOAA-GLERL contribution number XXXX.

Conflict of interest.

The authors declare they do not have any competing financial interests in relation to this work.

Figure Legends:

Figure 1: A) Map of sampling locations in western Lake Erie. B) Photosynthetic pigment, toxin, and relative abundance of *Cyanobacteria* reads across sites and sampling dates. M denotes a missing sample.

Figure 2: Cyanobacterial spatial and temporal dynamics during the western Lake Erie CHAB. A) Cyanobacterial genus composition across stations and timepoints. Relative abundance is measured with respect to the total bacterial community. B) Cyanobacterial OTU temporal dynamics. OTUs with mean relative abundance > 0.0001 are depicted. Relative abundance is measured with respect to the total bacterial community.

Figure 3: Nc-bacterial alpha diversity trends. A) Nc-bacterial observed richness trends over time. B-D) Observed richness of *Alphaproteobacteria*, *Bacteroidetes*, and *Betaproteobacteria* with respect to log chl *a* concentrations. E) Nc-bacterial evenness measured by Simpson's E over time. F-H) Evenness of *Alphaproteobacteria*, *Bacteroidetes*, and *Betaproteobacteria* with respect to chl *a* concentrations. Reported p-

values underwent FDR correction for multiple hypotheses. For plots of other bacterial groups and correlation to pH and phycocyanin see figures S2-S3.

Figure 4: Principal coordinates analyses of nc-bacterial Bray-Curtis dissimilarity. Three principal coordinates were selected based on the output of a scree plot.
A-B) PC 1 scores with respect to time and pH. C-D) PC 2 scores with respect to time and temperature. E-F) PC3 scores with respect to time and water specific conductivity.

Figure 5: Spatial and temporal dynamics of abundant *Actinobacteria* acI OTUs.

Table1: Regression models to predict scores on Bray-Curtis principal coordinates over time. The top model(s) for each PC are reported. Only one environmental covariate was considered in each model, and models were compared with and without time as an additional covariate. P-values underwent FDR correction. Cross validation was performed by leaving out all samples from the same timepoint as the test set.
Supplementary plots showing model residuals are in Figure S5-S6.

Variable	**PC1**	**PC1**	**PC2**	**PC3**
model	~ pH + time	~logChla + time	~ Temperature	~SpCond
p-value	***	***	***	***
R^2	0.678	0.658	0.822	0.451
Cross-validation MSE	0.01	0.0233	0.0522	0.0494

Literature.

Agha, R., del Mar Labrador, M., de los Ríos, A., and Quesada, A. (2016) Selectivity and detrimental effects of epiphytic Pseudanabaena on Microcystis colonies. *Hydrobiologia* **777**: 139–148.

Aitchison, J. (1982) The Statistical Analysis of Compositional Data. *J. R. Stat. Soc. Ser. B* **44**: 139–177.

Allgaier, M. and Grossart, H.-P. (2006) Diversity and Seasonal Dynamics of Actinobacteria Populations in Four Lakes in Northeastern Germany. *Appl. Environ. Microbiol.* **72**: 3489–3497.

Apprill, A., McNally, S., Parsons, R., and Weber, L. (2015) Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquat. Microb. Ecol.* **75**: 129–137.

Backer, L., Landsberg, J., Miller, M., Keel, K., and Taylor, T. (2013) Canine Cyanotoxin Poisonings in the United States (1920s–2012): Review of Suspected and Confirmed Cases from Three Data Sources. *Toxins (Basel).* **5**: 1597–1628.

Baho, D.L., Futter, M.N., Johnson, R.K., and Angeler, D.G. (2015) Assessing temporal scales and patterns in time series: Comparing methods based on redundancy analysis. *Ecol. Complex.* **22**: 162–168.

Behrenfeld, M.J., Boss, E., Siegel, D.A., and Shea, D.M. (2005) Carbon-based ocean productivity and phytoplankton physiology from space. *Global Biogeochem. Cycles* **19**: 1–14.

Bergmann, G.T., Bates, S.T., Eilers, K.G., Lauber, C.L., Caporaso, J.G., Walters, W.A., et al. (2011) The under-recognized dominance of Verrucomicrobia in soil bacterial

communities. *Soil Biol. Biochem.* **43**: 1450–1455.

Bingham, M., Sinha, S.K., and Lupi, F. (2015) Economic benefits of reducing harmful algal blooms in Lake Erie.

Bridgeman, T.B., Chaffin, J.D., and Filbrun, J.E. (2013) A novel method for tracking western Lake Erie Microcystis blooms, 2002–2011. *J. Great Lakes Res.* **39**: 83–89.

Brooks, J.P., Edwards, D.J., Harwich, M.D., Rivera, M.C., Fettweis, J.M., Serrano, M.G., et al. (2015) The truth about metagenomics: quantifying and counteracting bias in 16S rRNA studies. *BMC Microbiol.* **15**: 66.

Carmichael, W.W., Azevedo, S.M., An, J.S., Molica, R.J., Jochimsen, E.M., Lau, S., et al. (2001) Human fatalities from cyanobacteria: chemical and biological evidence for cyanotoxins. *Environ. Health Perspect.* **109**: 663–8.

Conroy, J.D., Edwards, W.J., Pontius, R.A., Kane, D.D., Zhang, H., Shea, J.F., et al. (2005) Soluble nitrogen and phosphorus excretion of exotic freshwater mussels (Dreissena spp.): Potential impacts for nutrient remineralisation in western Lake Erie. *Freshw. Biol.* **50**: 1146–1162.

Cory, R.M., Davis, T.W., Dick, G.J., Johengen, T., Denef, V.J., Berry, M., et al. (2016) Seasonal dynamics in dissolved organic matter, hydrogen peroxide, and cyanobacterial blooms in Lake Erie. *Front. Mar. Sci.* **3**: 54.

Cram, J.A., Chow, C.-E.T., Sachdeva, R., Needham, D.M., Parada, A.E., Steele, J.A., and Fuhrman, J.A. (2015) Seasonal and interannual variability of the marine bacterioplankton community throughout the water column over ten years. *Isme J* **9**: 563–580.

Crump, B.C. and Hobbie, J.E. (2005) Synchrony and seasonality in bacterioplankton

communities of two temperate rivers. *Limnol. Oceanogr.* **50**: 1718–1729.

Davis, T.W. and Gobler, C.J. (2016) Preface for special issue on global expansion of harmful cyanobacterial blooms: Diversity, ecology, causes, and controls. *Harmful Algae*.

Davis, T.W., Koch, F., Marcoval, M.A., Wilhelm, S.W., and Gobler, C.J. (2012) Mesozooplankton and microzooplankton grazing during cyanobacterial blooms in the western basin of Lake Erie. *Harmful Algae* **15**: 26–35.

Downing, A.L. and Leibold, M.A. (2002) Ecosystem consequences of species richness and composition in pond food webs. *Nature* **416**: 837–841.

Eiler, A. and Bertilsson, S. (2004) Composition of freshwater bacterial communities associated with cyanobacterial blooms in four Swedish lakes. *Environ. Microbiol.* **6**: 1228–1243.

Eiler, A., Heinrich, F., and Bertilsson, S. (2012) Coherent dynamics and association networks among lake bacterioplankton taxa. *ISME J.* **6**: 330–42.

Freeman, J., Vladimirov, N., Kawashima, T., Mu, Y., Sofroniew, N.J., Bennett, D. V., et al. (2014) Mapping brain activity at scale with cluster computing. *Nat. Methods* **11**: 941–950.

Gobler, C.J., Burkholder, J.M., Davis, T.W., Harke, M.J., Stow, C.A., and Van de Waal, D.B. (2016) The dual role of nitrogen supply in controlling the growth and toxicity of cyanobacterial blooms. *Harmful Algae*.

Gobler, C.J., Davis, T.W., Deonarine, S.N., Saxton, M.A., Lavrentyev, P.J., Jochem, F.J., and Wilhelm, S.W. (2008) Grazing and virus-induced mortality of microbial populations before and during the onset of annual hypoxia in Lake Erie. *Aquat.*

Microb. Ecol. **51**: 117–128.

Harke, M.J., Davis, T.W., Watson, S.B., and Gobler, C.J. (2016) Nutrient-Controlled Niche Differentiation of Western Lake Erie Cyanobacterial Populations Revealed via Metatranscriptomic Surveys. *Environ. Sci. Technol.* **50**: 604–15.

Harke, M.J., Steffen, M.M., Gobler, C.J., Otten, T.G., Wilhelm, S.W., Wood, S.A., and Paerl, H.W. (2016) A review of the global ecology, genomics, and biogeography of the toxic cyanobacterium, Microcystis spp. *Harmful Algae* **54**: 4–20.

Hernández, J.M., López-Rodas, V., and Costas, E. (2009) Microcystins from tap water could be a risk factor for liver and colorectal cancer: A risk intensified by global change. *Med. Hypotheses* **72**: 539–540.

Horner-Devine, M.C., Leibold, M. a, Smith, V.H., and Bohannan, B.J.M. (2003) Bacterial diversity patterns along a gradient of primary productivity. *Ecol. Lett.* **6**: 613–622.

Hu, Z., Liu, Y., and Li, D. (2004) Physiological and biochemical analyses of microcystin-RR toxicity to the cyanobacterium Synechococcus elongatus. *Environ. Toxicol.* **19**: 571–7.

Johansen, S. (2007) Correlation, regression, and cointegration of nonstationary economic time series. *Creat. Res. Pap.* **2461**: 0–9.

Jones, S.E., Newton, R.J., and McMahon, K.D. (2009) Evidence for structuring of bacterial community composition by organic carbon source in temperate lakes. *Environ. Microbiol.* **11**: 2463–2472.

Kara, E.L., Hanson, P.C., Hu, Y.H., Winslow, L., and McMahon, K.D. (2012) A decade of seasonal dynamics and co-occurrences within freshwater bacterioplankton

communities from eutrophic Lake Mendota, WI, USA. *ISME J.* **7**: 680–684.

Kent, A.D., Jones, S.E., Yannarell, A.C., Graham, J.M., Lauster, G.H., Kratz, T.K., and Triplett, E.W. (2004) Annual Patterns in Bacterioplankton Community Variability in a Humic Lake. *Microb. Ecol.* **48**: 550–560.

Lau, W.W.Y., Keil, R.G., and Armbrust, E. V. (2007) Succession and Diel Transcriptional Response of the Glycolate-Utilizing Component of the Bacterial Community during a Spring Phytoplankton Bloom. *Appl. Environ. Microbiol.* **73**: 2440–2450.

Lauber, C.L., Hamady, M., Knight, R., and 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.

Lindstrom, E.S., Kamst-Van Agterveld, M.P., and Zwart, G. (2005) Distribution of Typical Freshwater Bacterial Groups Is Associated with pH, Temperature, and Lake Water Retention Time. *Appl. Environ. Microbiol.* **71**: 8201–8206.

Lingoes, J.C. (1971) Some boundary conditions for a monotone analysis of symmetric matrices. *Psychometrika* **36**: 195–203.

Llirós, M., Inceoğlu, Ö., García-Armisen, T., Anzil, A., Leporcq, B., Pigneur, L.-M., et al. (2014) Bacterial Community Composition in Three Freshwater Reservoirs of Different Alkalinity and Trophic Status. *PLoS One* **9**: e116145.

López-Archilla, A.I., Moreira, D., López-García, P., and Guerrero, C. (2004) Phytoplankton diversity and cyanobacterial dominance in a hypereutrophic shallow lake with biologically produced alkaline pH. *Extremophiles* **8**: 109–115.

Maurice, C.F., CL Knowles, S., Ladau, J., Pollard, K.S., Fenton, A., Pedersen, A.B., and

Turnbaugh, P.J. (2015) Marked seasonal variation in the wild mouse gut microbiota. *ISME J.* **9**: 2423–2434.

McMurdie, P.J. and Holmes, S. (2013) phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* **8**: e61217.

McMurdie, P.J. and Holmes, S. (2014) Waste not, want not: why rarefying microbiome data is inadmissible. *PLoS Comput. Biol.* **10**: e1003531.

Michalak, A.M., Anderson, E.J., Beletsky, D., Boland, S., Bosch, N.S., Bridgeman, T.B., et al. (2013) Record-setting algal bloom in Lake Erie caused by agricultural and meteorological trends consistent with expected future conditions. *Proc. Natl. Acad. Sci.* **110**: 6448–6452.

Miller, M.A., Kudela, R.M., Mekebri, A., Crane, D., Oates, S.C., Tinker, M.T., et al. (2010) Evidence for a Novel Marine Harmful Algal Bloom: Cyanotoxin (Microcystin) Transfer from Land to Sea Otters. *PLoS One* **5**: e12576.

Millie, D.F., Fahnenstiel, G.L., Dyble, J., Ae, B., Pigg, R.J., Rediske, R.R., et al. Late-summer phytoplankton in western Lake Erie (Laurentian Great Lakes): bloom distributions, toxicity, and environmental influences.

Newton, R.J., Jones, S.E., Eiler, A., McMahon, K.D., and Bertilsson, S. (2011) A guide to the natural history of freshwater lake bacteria.

Newton, R.J., Jones, S.E., Helmus, M.R., and McMahon, K.D. (2007) Phylogenetic ecology of the freshwater Actinobacteria acI lineage. *Appl. Environ. Microbiol.* **73**: 7169–7176.

O'Neil, J.M., Davis, T.W., Burford, M.A., and Gobler, C.J. (2012) The rise of harmful cyanobacteria blooms: The potential roles of eutrophication and climate change.

Harmful Algae **14**: 313–334.

Ohio Department of Natural Resources, D. of G.S. Lake Erie Facts.

Ouellette, A.J.A., Handy, S.M., and Wilhelm, S.W. (2006) Toxic Microcystis is widespread in Lake Erie: PCR detection of toxin genes and molecular characterization of associated cyanobacterial communities. *Microb. Ecol.* **51**: 154–165.

Paerl, H.W. and Huisman, J. (2009) Climate change: a catalyst for global expansion of harmful cyanobacterial blooms. *Environ. Microbiol. Rep.* **1**: 27–37.

Paerl, H.W. and Otten, T.G. (2013) Harmful Cyanobacterial Blooms: Causes, Consequences, and Controls. *Microb. Ecol.* **65**: 995–1010.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013) The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **41.**:

R Core Team (2015) R: a language for statistical computing.

Rangel, M., Martins, J., Garcia, A., Conserva, G., Costa-Neves, A., Sant'Anna, C., and de Carvalho, L. (2014) Analysis of the Toxicity and Histopathology Induced by the Oral Administration of Pseudoanabaena galeata and Geitlerinema splendidum (Cyanobacteria) Extracts to Mice. *Mar. Drugs* **12**: 508–524.

Revelle, W. (2015) psych: Procedures for personality and psychological Research.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., et al. (2009) Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. *Appl. Environ. Microbiol.* **75**: 7537–7541.

Shade, A., Kent, A.D., Jones, S.E., Newton, R.J., Triplett, E.W., and McMahon, K.D. (2007) Interannual dynamics and phenology of bacterial communities in a eutrophic lake. *Limnol. Oceanogr.* **52**: 487–494.

Shade, A., Peter, H., Allison, S.D., Baho, D.L., Berga, M., B??rgmann, H., et al. (2012) Fundamentals of microbial community resistance and resilience. *Front. Microbiol.* **3**: 1–19.

Shade, A., Read, J.S., Welkie, D.G., Kratz, T.K., Wu, C.H., and McMahon, K.D. (2011) Resistance, resilience and recovery: Aquatic bacterial dynamics after water column disturbance. *Environ. Microbiol.* **13**: 2752–2767.

Smith, V.H. (2007) Microbial diversity-productivity relationships in aquatic ecosystems. *FEMS Microbiol. Ecol.* **62**: 181–186.

Steffen, M.M., Belisle, B.S., Watson, S.B., Boyer, G.L., and Wilhelm, S.W. (2014) Status, causes and controls of cyanobacterial blooms in Lake Erie. *J. Great Lakes Res.* **40**: 215–225.

Tanber, G. (2014) Toxin leaves 500,000 in northwest Ohio without drinking water. *Reuters*.

Tang, X., Gao, G., Chao, J., Wang, X., Zhu, G., and Qin, B. (2010) Dynamics of organic-aggregate-associated bacterial communities and related environmental factors in Lake Taihu, a large eutrophic shallow lake in China. *Limnol. Oceanogr.* **55**: 469–480.

Teeling, H., Fuchs, B.M., Becher, D., Klockow, C., Gardebrecht, A., Bennke, C.M., et al. (2012) Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science* **336**: 608–11.

U.S. EPA (1979) Methods for Chemical Analysis of Water and Wastes Cincinnati, OH.

Visser, P.M., Verspagen, J.M.H., Sandrini, G., Stal, L.J., Matthijs, H.C.P., Davis, T.W., et al. (2016) How rising CO2 and global warming may stimulate harmful cyanobacterial blooms. *Harmful Algae* **accepted.**:

Wang, Q., Garrity, G.M., Tiedje, J.M., and Cole, J.R. (2007) Naive Bayesian Classifier for Rapid Assignment of rRNA Sequences into the New Bacterial Taxonomy. *Appl. Environ. Microbiol.* **73**: 5261–5267.

Wilhelm, S.W., Farnsley, S.E., LeCleir, G.R., Layton, A.C., Satchwell, M.F., DeBruyn, J.M., et al. (2011) The relationships between nutrients, cyanobacterial toxins and the microbial community in Taihu (Lake Tai), China. *Harmful Algae* **10**: 207–215.

Woodhouse, J.N., Kinsela, A.S., Collins, R.N., Bowling, L.C., Honeyman, G.L., Holliday, J.K., and Neilan, B.A. (2015) Microbial communities reflect temporal changes in cyanobacterial composition in a shallow ephemeral freshwater lake. *ISME J.* 1–15.

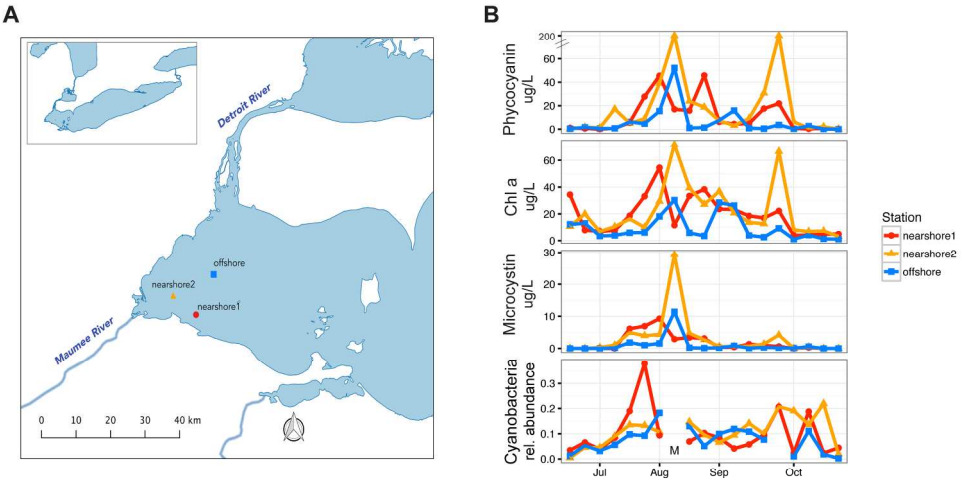
Xiong, J., Liu, Y., Lin, X., Zhang, H., Zeng, J., Hou, J., et al. (2012) Geographic distance and pH drive bacterial distribution in alkaline lake sediments across Tibetan Plateau. *Environ. Microbiol.* **14**: 2457–2466.

Yang, C., Li, Y., Zhou, B., Zhou, Y., Zheng, W., Tian, Y., et al. (2015) Illumina sequencing-based analysis of free-living bacterial community dynamics during an Akashiwo sanguine bloom in Xiamen sea, China. *Sci. Rep.* **5**: 8476.

Ye, W., Tan, J., Liu, X., Lin, S., Pan, J., Li, D., and Yang, H. (2011) Temporal variability of cyanobacterial populations in the water and sediment samples of Lake Taihu as determined by DGGE and real-time PCR. *Harmful Algae* **10**: 472–479.

Table 1: Regression models to predict scores on Bray-Curtis principal coordinates over time. The top model(s) for each PC are reported. Only one environmental covariate was considered in each model, and models were compared with and without time as an additional covariate. P-values underwent FDR correction. Cross validation was performed by leaving out all samples from the same timepoint as the test set. Supplementary plots showing model residuals are in Figure S5. MSE = mean squared error.

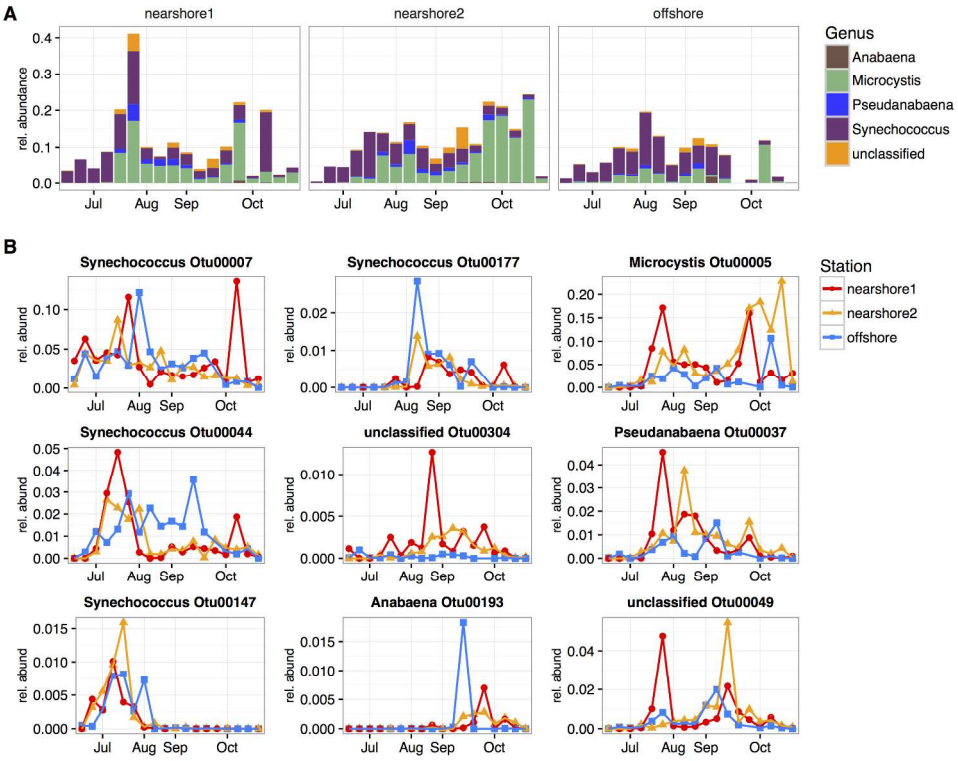
Variable	PC1	PC1	PC2	PC3
model	~ pH + time	~logChla + time	~ Temperature	~SpCond
p-value	<0.001	<0.001	<0.001	<0.001
R^2	0.678	0.658	0.822	0.451
Cross-validation MSE	0.01	0.0233	0.0522	0.0494



Sample sites and bloom dynamics. (A) Map of sampling locations in western Lake Erie. (B) Photosynthetic pigment, toxin, and relative abundance of Cyanobacteria reads across sites and sampling dates. M denotes a missing sample.

Figure 1

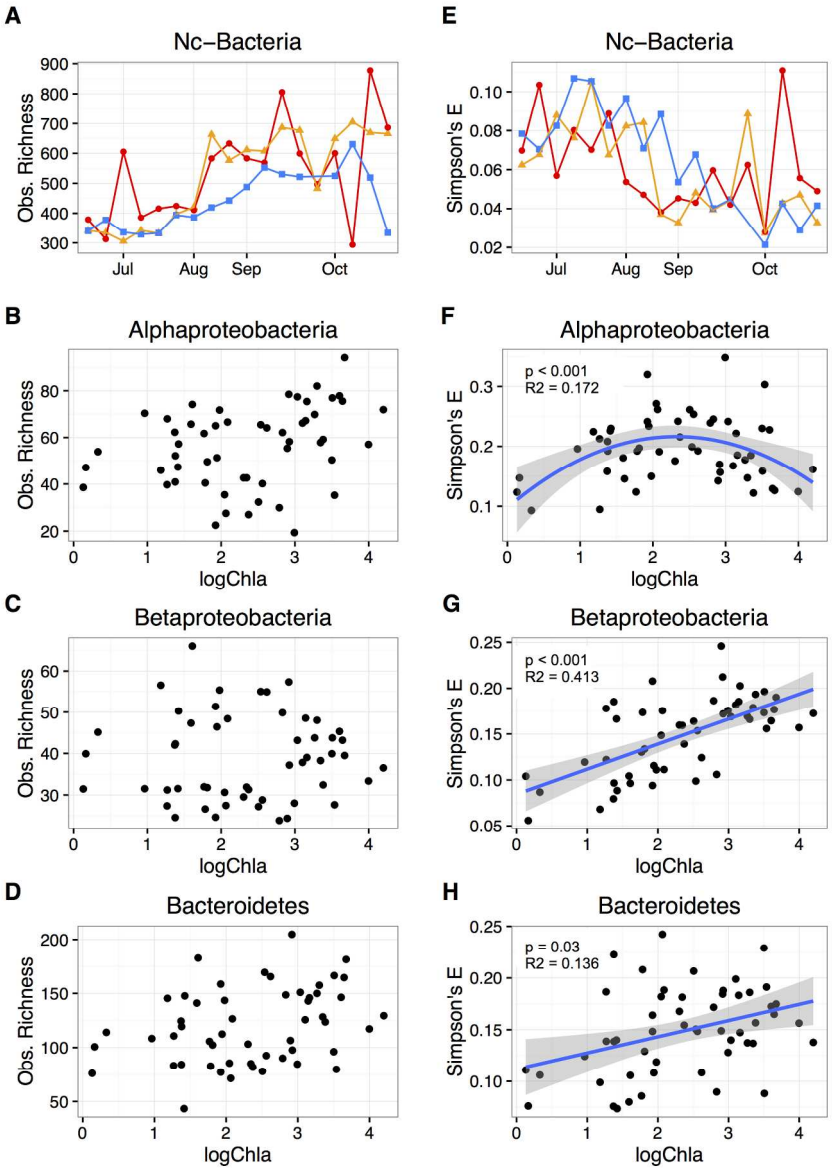
216x185mm (300 x 300 DPI)



Cyanobacterial spatial and temporal dynamics during the western Lake Erie CHAB. (A) Cyanobacterial genus composition across stations and timepoints. Relative abundance is measured with respect to the total bacterial community. (B) Cyanobacterial OTU temporal dynamics. OTUs with mean relative abundance > 0.0001 are depicted. Relative abundance is measured with respect to the total bacterial community.

Figure 2

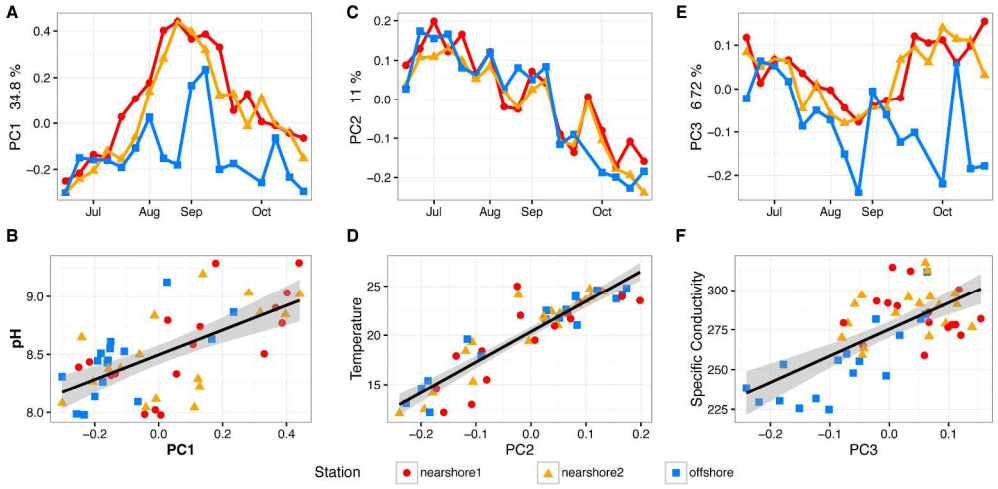
254x203mm (300 x 300 DPI)



Nc-bacterial alpha diversity trends. (A) Nc-bacterial observed richness trends over time. (B-D) Observed richness of Alphaproteobacteria, Bacteroidetes, and Betaproteobacteria with respect to log chl a concentrations. (E) Nc-bacterial evenness measured by Simpson's E over time. (F-H) Evenness of Alphaproteobacteria, Bacteroidetes, and Betaproteobacteria with respect to chl a concentrations. Reported p-values underwent FDR correction for multiple hypotheses. For plots of other bacterial groups and correlation to pH and phycocyanin see figures S2-S3.

Figure 3

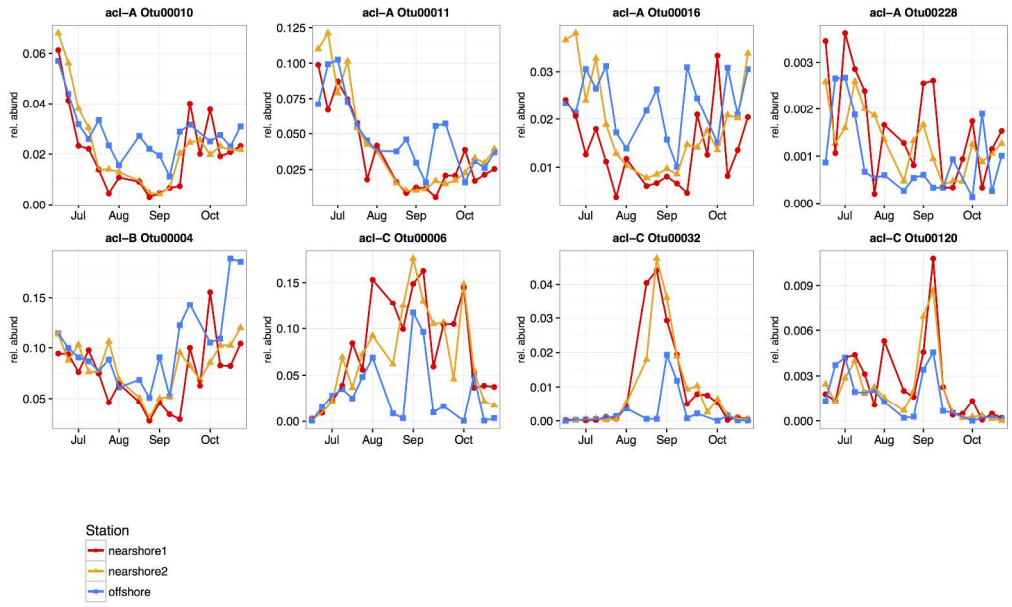
177x254mm (300 x 300 DPI)



Principal coordinates analyses of nc-bacterial Bray-Curtis dissimilarity. Three principal coordinates were selected based on the output of a scree plot (Figure S4).‼ ┼ (A-B) PC 1 scores with respect to time and pH. (C-D) PC 2 scores with respect to time and temperature. (E-F) PC3 scores with respect to time and water specific conductivity.‼ ┼

Figure 4

262x126mm (300 x 300 DPI)



Spatial and temporal dynamics of abundant Actinobacteria acI OTUs.
Figure 5
300x179mm (300 x 300 DPI)