Are oligotypes meaningful ecological and phylogenetic units? A case study of *Microcystis* in freshwater lakes

Michelle A. Berry[1], Jeffrey D. White[2], Timothy W. Davis[3], Sunit Jain[4,§], Thomas H. Johengen[5], Gregory J. Dick[4], Orlando Sarnelle[6], and Vincent J. Denef[1*]

[1]Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI, 48109.

[2]Department of Biology, Framingham State University, Framingham, MA, 01702.

[3]NOAA Great Lakes Environmental Research Laboratory, Ann Arbor MI, 48108.

[4]Department of Earth and Environmental Sciences, University of Michigan, Ann Arbor, MI, 48109

[5]Cooperative Institute for Limnology and Ecosystems Research, University of Michigan, Ann Arbor, MI, 48109.

[6]Department of Fisheries and Wildlife, Michigan State University, East Lansing, MI, 48824

*** Correspondence:**
Dr. Vincent Denef
vdenef@umich.edu

Keywords: oligotypes, microbial species, ecotypes, dada2, 16S rRNA gene sequencing, *Microcystis*

Abstract

Oligotyping is a computational method aimed at increasing the resolution of operational taxonomic units in 16S rRNA gene amplicon-based microbiome studies. Although oligotyping can distinguish highly similar sequence variants, the resulting units are not necessarily phylogenetically and ecologically informative. We critically examine claims from the recent oligotyping literature, and we illustrate the method's limitations with a case study of the harmful bloom-forming cyanobacterium *Microcystis*. We identified three *Microcystis* oligotypes from a western Lake Erie bacterial community 16S rRNA gene (V4 region) survey that had previously clustered into one OTU. We found the same three oligotypes and two additional sequence variants in 46 *Microcystis* cultures isolated from Michigan inland lakes spanning a trophic gradient. In Lake Erie, shifts in *Microcystis* oligotypes corresponded to spatial nutrient gradients and temporal transitions in bloom toxicity. In the cultures, *Microcystis* oligotypes showed preferential distributions for different trophic habitats, but genomic data revealed that the oligotypes identified in Lake Erie did not correspond to toxin gene presence. Thus, oligotypes could not be used for inferring toxic ecotypes. Most strikingly, *Microcystis* oligotypes were not monophyletic. Our study supports the utility of oligotyping for distinguishing populations along certain ecological features, while it highlights that such populations may not be ecologically or phylogenetically cohesive. Therefore, studies employing oligotyping or related tools must critically consider these caveats during data interpretation.

Main

1.1 Precarious interpretations of oligotypes in the microbial ecology literature

Microbiome studies using 16S rRNA gene amplicons typically aggregate sequences into operational taxonomic units (OTUs) based on a 97% sequence similarity threshold. The OTU approach is used, in part, to mitigate effects of high error rates from high-throughput sequencing technologies. However, OTU methods "throw out" potentially informative 16S sequence variation and can group together ecologically distinct populations (Coleman et al., 2006; Denef et al., 2010; Hunt et al., 2008; Shapiro and Polz, 2014).

As an alternative to OTUs, the oligotyping method is designed to distinguish real sequence variants from sequencing errors, and can segregate sequence types based on a single nucleotide difference (Eren et al., 2013). This increased resolution offered by this approach allegedly enhances the likelihood of identifying ecologically meaningful units. Since oligotyping was published, it has been used to investigate questions in microbial biogeography (Buttigieg and Ramette, 2015; Cloutier et al., 2015; Newton and McLellan, 2015; Schmidt et al., 2014), host-microbe associations (Eren et al., 2014; Fisher et al., 2015; Menke et al., 2014), and links between microbes and disease (Eren et al., 2011). Throughout this growing literature, there are pervasive assumptions about the ecological and phylogenetic cohesiveness of oligotypes, which have not been properly validated.

A common assertion from oligotyping studies is that 16S rRNA amplicon sequence variants represent distinct ecological units, ecotypes, or species-like groups (Delmont et al., 2014; Kleindienst et al., 2015; Schmidt et al., 2014)(Delmont et al., 2014; Kleindienst et al., 2015; Schmidt et al., 2014). For example, in a study of an oil well blowout in the Gulf of Mexico, the authors state, "Oligotyping provided a way to quantify, with great precision, the response of microbial taxa likely representing distinct ecotypes during this acute environmental perturbation" (Kleindienst et al., 2015). Although the meaning of a microbial species or ecotype is still highly debated, experimental and theoretical work has converged on a definition that includes inhabiting the same ecological niche,

exhibiting constrained genetic diversity, and belonging to a distinct evolutionary lineage (Cohan and Perry, 2007; Gevers et al., 2005; Koeppel et al., 2008). Therefore, papers that link oligotypes to ecotypes make general assumptions about the ecological, genetic, and evolutionary cohesiveness of populations defined by fine-scale nucleotide variation in 16S hypervariable regions. In contradiction, there are several lines of evidence demonstrating that the 16S rRNA gene, even at full length, can miss important genetic variation underlying ecological and evolutionary differentiation between species (Hahn et al., 2016; Jaspers and Overmann, 2004; Kim et al., 2014; Konstantinidis and Tiedje, 2005; Maiden et al., 2013). From an ecological perspective, many bacterial functional traits are not phylogenetically conserved and therefore are unlikely to be predicted from the 16S rRNA gene (Martiny et al., 2013). From an evolutionary perspective, the 16S rRNA gene is a slow evolving gene (Ochman et al., 1999), that while useful for assigning higher level bacterial taxonomy, may not encode more recent evolutionary diversification within a lineage. Thus, the 16S rRNA gene has limited sensitivity, and all oligotyping analyses aimed at inferring ecotypes are inherently constrained by this attribute.

Another precarious use of oligotyping in the literature is for supporting co-evolutionary hypotheses between hosts and microbes. A paper on the gut microbiomes of sympatric Namibian carnivore species used oligotyping to distinguish bacterial groups found in jackals and cheetahs that could not be resolved at the OTU level. The authors suggest that differences between these two species might be due to a "co-evolutionary fine-tuning of some genera according to the digestive requirements of the host" (Menke et al., 2014). This statement, like claims about ecotypes, suggests that oligotypes are phylogenetically distinct lineages upon which co-evolution would act.

Finally, oligotyping has been frequently used to identify microbial biogeographic patterns (Buttigieg and Ramette, 2015; Newton and McLellan, 2015; Schmidt et al., 2014), which are subject to similar ecological and evolutionary assumptions as ecotype inferences. In a comparative study of bacterial communities found across Lake Michigan beaches in Michigan and Wisconsin, the authors write, "within predominate genera, fine-scale sequence differences could be found that distinguished the populations from the two states, suggesting a biogeographic effect" (Cloutier et al., 2015). The goal of a biogeographic study is to find biological and geographic patterns to explain the distribution of species populations. If oligotypes do not, in fact, represent cohesive ecological or phylogenetic populations, the meaning that can be derived from an oligotyping biogeographic study is greatly reduced.

In addition to oligotyping, a tool called dada2 was recently developed to resolve sub-OTU amplicon sequence variants (Callahan et al., 2016). The algorithmic approach differs between dada2 and oligotyping; oligotyping uses Shannon's entropy, while dada2 explicitly models transition rates between closely related sequences. However, both methods are aimed at identifying ecologically refined populations, and we expect similar adoption. Therefore, the concerns we raise with respect to oligotyping also apply to dada2 and any other tools attempting to make ecotype inferences from 16S rRNA gene amplicons.

1.2 Microcystis case study

To illustrate potential issues with the ecological and evolutionary assumptions made about oligotypes in the aforementioned literature, we present a case-study of *Microcystis*, a colony-forming cyanobacterium that is a prominent component of harmful algal blooms in freshwater systems worldwide (Harke et al., 2016b; O'Neil et al., 2012). First, we oligotyped *Microcystis* reads from a bacterial community dataset sampled over three sites and twenty weeks from western Lake Erie

during the 2014 cyanobacterial harmful algal bloom. We analyzed oligotypes with respect to two key ecological traits: toxic potential and trophic preference. This approach is comparable to the setup of the articles examined above, in that we attempted to link uncharacterized oligotypes from a community dataset to environmental gradients. Next, we examined oligotypes and genomes from 46 non-axenic *Microcystis* cultures, which were isolated from 14 Michigan inland lakes in 2011 and 2013. Comparing oligotypes from the Lake Erie community samples with the cultures provided multiple advantages. The cultures were isolated from single *Microcystis* colonies and were typically of clonal origin, which served to constrain the considered population. In addition, the culture collection allowed us to compare, with high accuracy, the gene content of each *Microcystis* isolate with its oligotype, and to construct a Multi-Locus Sequence Typing (MLST) phylogeny based on housekeeping genes that would be difficult to recover from a community dataset due to coverage or binning issues.

Microcystis was the dominant large-colony forming cyanobacterial genus in the 2014 cyanobacterial harmful algal bloom in western Lake Erie. Analysis with mothur produced a single abundant *Microcystis* OTU, while oligotyping produced subdivisions of the OTU into three sequence variants (CTG, CCG, CTT), which exhibited differing spatial and temporal dynamics (Figure 1). First, we observed that the CTG variant dominated in July and August (median CTG:CCG ratio = 4.3), but the CCG variant dominated in September and October (median CTG:CCG ratio = 0.23). The transition between these two sequence variants coincided with a shift in bloom toxicity from high to low (Figure 1), a trend that has been documented in other bloom years on Lake Erie (Gobler et al., 2016). We hypothesized that CTG might represent a toxic ecotype, and CCG might represent a non-toxic ecotype. Indeed, the relative abundance of CTG was positively correlated with particulate microcystin-LR levels (Spearman's rho: 0.71, $p < 0.001$). Second, we observed that the median relative abundance of CTT was an order of magnitude higher at the offshore station than the nearshore stations (permutational test for significantly different medians: $p = 0.0007$). Since the offshore station had lower median phosphorus levels (Table S1), we hypothesized that the oligotypes might underlie differences in competitive abilities along trophic gradients.

Next, we examined 16S rRNA gene and whole genome data from a collection of *Microcystis* isolate cultures (Table S2) to further investigate our hypotheses about *Microcystis* oligotypes, toxicity, and trophic status. Similar to the Lake Erie dataset, all *Microcystis* 16S rRNA gene V4 region sequences clustered into one OTU, but we recovered five oligotypes (CTG, CCG, CTT, TCG, CCT). Three of these matched the oligotypes found in Lake Erie. Although the oligotypes derived from the cultures did not indicate the trophic status of the inland lakes, the trophic status could predict which oligotypes were present (Figure 2). For example, CTT was the only oligotype present in oligotrophic lakes, CCG and CTG were the only oligotypes present in mesotrophic lakes, but all five oligotypes were present in inland eutrophic lakes. These data support that fine-scale variation in the 16S rRNA gene V4 region might distinguish populations with differing competitive abilities along nutrient gradients. Specifically, CTT might exclude other oligotypes from oligotrophic environments, yet all three oligotypes might coexist in the eutrophic environments due to intra-lake spatial or temporal variation in nutrient concentrations. Other recent work from western Lake Erie demonstrates that *Microcystis* populations upregulate phosphorus scavenging genes in response to low phosphorus conditions at offshore sites, leading to a competitive advantage over other cyanobacterial taxa (Harke et al., 2016a). Our data suggests that low phosphorus conditions could also select for particular *Microcystis* oligotypes, a hypothesis that will need to be formally tested.

A second hypothesis derived from our Lake Erie observations was that *Microcystis* oligotypes represent ecotypes that differ in their ability to produce toxins. However, when examining the

cultures, it became apparent that the oligotypes did not unequivocally correspond to the presence of genes for microcystin biosynthesis. These data are consistent with previous reports that strains containing the toxin producing *mcy* gene cluster form a polyphyletic group in *Microcystis* and other toxin-producing *Cyanobacteria* (Kurmayer et al., 2014; Otsuka et al., 1999). Therefore, despite a correlation between oligotypes and toxicity in the Lake Erie dataset, we could not corroborate that the CTG variant represents a toxic ecotype and the CCG variant represents a nontoxic ecotype. Furthermore, a recent review of global *Microcystis* diversity indicates that 27 strains exhibit 99.4-99.93% similarity across the full length 16S rRNA gene, so surveys based on shortened amplicons are likely to group several mixed populations together (Harke et al., 2016b).

As for the evolutionary interpretation of oligotypes, an MLST analysis performed on the culture genomic data revealed two main results. First, strains of the same oligotype did not always form monophyletic groups (Figure 2). Second, the number of nucleotide differences in the 16S rRNA gene V4 region was not consistent with MLST-based patristic distances (Figure S2). For example, oligotypes with two or three nucleotide differences in the 16S V4 region were not more distant on the MLST tree than oligotypes differing by one nucleotide. This indicated that *Microcystis* V4 oligotypes are not phylogenetically cohesive units. Previous work has shown that, surprisingly, many OTUs across bacterial phyla are not monophyletic(Koeppel and Wu, 2013). Our data support this observation and demonstrate the principle that 16S rRNA gene hypervariable regions can be poor proxies for evolutionary distance irrespective of what resolution is achieved.

In summary, distributions of *Microcystis* oligotypes from an environmental community dataset corresponded with shifts in toxicity and spatial variation in phosphorus levels. However, an additional analysis leveraging genomic data from *Microcystis* cultures revealed that oligotypes did not predict presence of toxin genes and the oligotypes were not monophyletic. This case study supports the idea that 16S rRNA gene amplicons may be useful to discriminate ecologically distinct populations when more complex and presumably multi-genic traits are considered, such as shifts on the oligotroph-copiotroph spectrum (Lauro et al., 2009; Martiny et al., 2013). However, when we focus on ecological traits that are underpinned by a single or a handful of genes in the flexible genome, such as toxin production, 16S rRNA gene amplicons may carry limited information and single nucleotide variants may lead us to unwarranted conclusions.

1.3 A call for a more critical interpretation of oligotyping and dada2 data

The observations from this *Microcystis* case study demonstrate that caution must be applied when using fine-scale sequence variation in 16S rRNA gene amplicons to make inferences about the ecology or phylogeny of uncharacterized populations in environmental samples. Returning to the studies referenced earlier, claims of sub-OTU sequence variants revealing distinct ecological units require further validation either through culture-based or metagenomic approaches. This validation is particularly important when an ecological classification for a sequence variant is made on the basis of one particular trait or environmental gradient. Kleindienst et al. (2015) claim “oligotypes that correlate significantly with environmental parameters likely represent distinct ecotypes”. However, our case study demonstrates that even when such associations are found, variation in the 16S rRNA gene might not correspond to variation in the trait being considered, and the resulting oligotypes are not necessarily ecologically cohesive. Therefore, it is imprudent to assert that sequence variants obtained from a shortened 16S amplicon represent ecologically distinct groups.

Our case study also demonstrates that assumptions about the phylogenetic cohesiveness of oligotypes may not be accurate, which makes equating oligotypes to populations or ecotypes highly problematic. Hypervariable regions of the 16S rRNA gene generally correlate with variation across the whole gene (Kim et al., 2011), but this correspondence is not perfect, so variation at a single nucleotide position may or may not be phylogenetically informative. Furthermore, the 16S rRNA gene is highly conserved across all domains of life, and while it has proven successful at resolving the more basal elements of trees, multiple housekeeping loci are considered much more precise for resolving relationships between taxa of the same genus (Maiden et al., 2013).

Oligotyping and dada2 are useful tools because they provide more *potential* biologically relevant information from high-throughput sequencing studies than OTU analyses. However, even at its maximum resolution, i.e. full-length high quality sequences, the 16S rRNA gene does not encode all ecologically or evolutionarily meaningful variation. Therefore, the increased resolution enabled by these new methods should not lure us away from rigorous use of ecological and evolutionary terms and concepts. We propose that oligotyping and dada2 should be used as a foundation to generate ecological hypotheses from microbial community datasets, but our case study reinforces that specific ecological and evolutionary patterns require validation by in-depth studies at the genomic level.

Figure Legends

Figure 1: Spatiotemporal distribution of *Microcystis* oligotypes in western Lake Erie. A) The relative abundance of *Microcystis* oligotypes, with respect to total bacterial reads, from three sites in Western Lake Erie over time. The offshore site had lower median total phosphorus and chlorophyll a levels than the two nearshore sites. Samples were taken from the retentate of 2 L lake water filtered through a 100 m filter. M denotes missing samples. B) Particulate Microcystin-LR concentrations over sites and time.

Figure 2: RAxML tree for cultured *Microcystis* strains based on five concatenated housekeeping genes (pgi, gltX, ftsZ, glnA, gyrB). Presence of microcystin biosynthesis gene was determined from assembly and retrieval of *Microcystis* genes from the non-axenic culture metagenome. Trophic status of the lake was determined from total phosphorus levels (Supplementary table 1 and 2).

Data Access

Sequences used for oligotyping analyses will be available under SRA accession numbers XXX. Sequences used for MLST analyses will be available under GenBank accession numbers XXX. All sample data and code to fully reproduce analyses will become available at https://github.com/DenefLab/Microcystis-Oligotypes

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author Contributions

MAB, JW, OS, TWD, THJ, GJD, VJD designed the experiment. JW, MAB, TWD, THJ collected data. MAB, SJ, VJD analyzed data. MAB, VJD wrote the paper.

Funding

GJD, THJ, TWD, and VJD were supported by a grant from the Erb Family Foundation made through the University of Michigan Water Center and by funds from EPA Great Lakes Restoration Initiative. JW and OS were supported by funds from the Gull Lake Quality Organization, and the Robert C. Ball and Betty A. Ball Fisheries and Wildlife Fellowship at Michigan State University.

Acknowledgments

We are grateful to the crew from NOAA-GLERL including Duane Gossiaux who assisted with fieldwork on Lake Erie. We thank C. Kozel, T. Geelhoed, and M. Schmidt for assistance sampling the inland lakes, M. Williams, A. Fogiel, and T. Brownell for providing logistical support, and C. Kozel, T. Geelhoed, and S. Flemming for maintaining the *Microcystis* cultures. This manuscript is NOAA-GLERL contribution number XXXX.

References

Buttigieg, P. L., and Ramette, A. (2015). Biogeographic patterns of bacterial microdiversity in Arctic deep-sea sediments (HAUSGARTEN, Fram Strait). *Front. Microbiol.* 6, 1–12. doi:10.3389/fmicb.2014.00660.

Callahan, B. J., Mcmurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. (2016). DADA2 : High-resolution sample inference from Illumina amplicon data. *Nat. Methods*, 1–7. doi:10.1038/nmeth.3869.

Cloutier, D. D., Alm, E. W., and McLellan, S. L. (2015). Influence of land use, nutrients, and geography on microbial communities and fecal indicator abundance at Lake Michigan beaches. *Appl. Environ. Microbiol.* 81, 4904–4913. doi:10.1128/AEM.00233-15.

Cohan, F. M., and Perry, E. B. (2007). A Systematics for Discovering the Fundamental Units of Bacterial Diversity. *Curr. Biol.* 17, R373–R386. doi:10.1016/j.cub.2007.03.032.

Coleman, M. L., Sullivan, M. B., Martiny, A. C., Steglich, C., Barry, K., Delong, E. F., et al. (2006). Genomic islands and the ecology and evolution of Prochlorococcus. *Science* 311, 1768–1770. doi:10.1126/science.1122050.

Delmont, T. O., Hammar, K. M., Ducklow, H. W., Yager, P. L., and Post, A. F. (2014). Phaeocystis antarctica blooms strongly influence bacterial community structures in the Amundsen Sea polynya. *Front. Microbiol.* 5, 646. doi:10.3389/fmicb.2014.00646.

Denef, V. J., Kalnejais, L. H., Mueller, R. S., Wilmes, P., Baker, B. J., Thomas, B. C., et al. (2010). Proteogenomic basis for ecological divergence of closely related bacteria in natural acidophilic microbial communities. *Proc. Natl. Acad. Sci.* 107, 2383–2390. doi:10.1073/pnas.0907041107.

Eren, a M., Borisy, G. G., Huse, S. M., and Mark Welch, J. L. (2014). Oligotyping analysis of the human oral microbiome. *Proc. Natl. Acad. Sci.* 111, E2875–E2884. doi:10.1073/pnas.1409644111.

Eren, A. M., Maignien, L., Sul, W. J., Murphy, L. G., Grim, S. L., Morrison, H. G., et al. (2013). Oligotyping: Differentiating between closely related microbial taxa using 16S rRNA gene data. *Methods Ecol. Evol.* 4, 1111–1119. doi:10.1111/2041-210X.12114.

Eren, A. M., Zozaya, M., Taylor, C. M., Dowd, S. E., Martin, D. H., and Ferris, M. J. (2011). Exploring the Diversity of Gardnerella vaginalis in the Genitourinary Tract Microbiota of Monogamous Couples Through Subtle Nucleotide Variation. *PLoS One* 6, e26732. doi:10.1371/journal.pone.0026732.

Fisher, J. C., Murat Eren, A., Green, H. C., Shanks, O. C., Morrison, H. G., Vineis, J. H., et al. (2015). Comparison of sewage and animal fecal microbiomes by using oligotyping reveals potential human fecal indicators in multiple taxonomic groups. *Appl. Environ. Microbiol.* 81, 7023–7033. doi:10.1128/AEM.01524-15.

Gevers, D., Cohan, F. M., Lawrence, J. G., Spratt, B. G., Coenye, T., Feil, E. J., et al. (2005). Opinion: Re-evaluating prokaryotic species. *Nat. Rev. Microbiol.* 3, 733–739. doi:10.1038/nrmicro1236.

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*.

Hahn, M. W., Jezberová, J., Koll, U., Saueressig-Beck, T., and Schmidt, J. (2016). Complete ecological isolation and cryptic diversity in Polynucleobacter bacteria not resolved by 16S rRNA gene sequences. *ISME J.* 10, 1642–1655. doi:10.1038/ismej.2015.237.

Harke, M. J., Davis, T. W., Watson, S. B., and Gobler, C. J. (2016a). Nutrient-Controlled Niche Differentiation of Western Lake Erie Cyanobacterial Populations Revealed via Metatranscriptomic Surveys. *Environ. Sci. Technol.* 50, 604–15. doi:10.1021/acs.est.5b03931.

Harke, M. J., Steffen, M. M., Otten, T. G., Wilhelm, S. W., Wood, S. A., and Paerl, H. W. (2016b). A review of the global ecology, genomics, and biogeography of the toxic cyanobacterium, *Microcystis* spp. *Harmful Algae* 54, 4–20. doi:10.1016/j.hal.2015.12.007.

Hunt, D. E., David, L. A., Gevers, D., Preheim, S. P., Alm, E. J., and Polz, M. F. (2008). Resource Partitioning and Sympatric Differentiation Among Closely Related Bacterioplankton. *Science (80-.).* 320, 1081–1085. doi:10.1126/science.1157890.

Jaspers, E., and Overmann, J. (2004). Ecological significance of microdiversity: identical 16S rRNA gene sequences can be found in bacteria with highly divergent genomes and ecophysiologies. *Appl. Environ. Microbiol.* 70, 4831–9. doi:10.1128/AEM.70.8.4831-4839.2004.

Kim, M., Morrison, M., and Yu, Z. (2011). Evaluation of different partial 16S rRNA gene sequence regions for phylogenetic analysis of microbiomes. *J. Microbiol. Methods* 84, 81–87. doi:10.1016/j.mimet.2010.10.020.

Kim, M., Oh, H.-S., Park, S.-C., and Chun, J. (2014). Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *Int. J. Syst. Evol. Microbiol.* 64, 346–51. doi:10.1099/ijs.0.059774-0.

Kleindienst, S., Grim, S., Sogin, M., Bracco, A., Crespo-Medina, M., and Joye, S. B. (2015). Diverse, rare microbial taxa responded to the Deepwater Horizon deep-sea hydrocarbon plume. *ISME J.* 10, 1–16. doi:10.1038/ismej.2015.121.

Koeppel, A. F., and Wu, M. (2013). Surprisingly extensive mixed phylogenetic and ecological signals among bacterial Operational Taxonomic Units. *Nucleic Acids Res.* 41, 5175–88. doi:10.1093/nar/gkt241.

Koeppel, A., Perry, E. B., Sikorski, J., Krizanc, D., Warner, A., Ward, D. M., et al. (2008). Identifying the fundamental units of bacterial diversity: a paradigm shift to incorporate ecology into bacterial systematics. *Proc. Natl. Acad. Sci. U. S. A.* 105, 2504–9. doi:10.1073/pnas.0712205105.

Konstantinidis, K. T., and Tiedje, J. M. (2005). Genomic insights that advance the species definition for prokaryotes. *Proc. Natl. Acad. Sci.* 102, 2567–2572. doi:10.1073/pnas.0409727102.

Kurmayer, R., Blom, J. F., Deng, L., and Pernthaler, J. (2014). Integrating phylogeny, geographic niche partitioning and secondary metabolite synthesis in bloom-forming Planktothrix. *ISME J.* 9, 909–921. doi:10.1038/ismej.2014.189.

Lauro, F. M., McDougald, D., Thomas, T., Williams, T. J., Egan, S., Rice, S., et al. (2009). The genomic basis of trophic strategy in marine bacteria. *Proc. Natl. Acad. Sci.* 106, 15527–15533. doi:10.1073/pnas.0903507106.

Maiden, M. C. J., Jansen van Rensburg, M. J., Bray, J. E., Earle, S. G., Ford, S. A., Jolley, K. A., et al. (2013). MLST revisited: the gene-by-gene approach to bacterial genomics. *Nat. Rev. Microbiol.* 11, 728–36. doi:10.1038/nrmicro3093.

Martiny, A. C., Treseder, K., and Pusch, G. (2013). Phylogenetic conservatism of functional traits in microorganisms. *ISME J.* 7, 830–838. doi:10.1038/ismej.2012.160.

Menke, S., Wasimuddin, Meier, M., Melzheimer, J., Mfune, J. K. E., Heinrich, S., et al. (2014). Oligotyping reveals differences between gut microbiomes of free-ranging sympatric Namibian carnivores (Acinonyx jubatus, Canis mesomelas) on a bacterial species-like level. *Front. Microbiol.* 5, 1–12. doi:10.3389/fmicb.2014.00526.

Newton, R. J., and McLellan, S. L. (2015). A unique assemblage of cosmopolitan freshwater bacteria and higher community diversity differentiate an urbanized estuary from oligotrophic Lake Michigan. *Front. Microbiol.* 6, 1–13. doi:10.3389/fmicb.2015.01028.

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. doi:10.1016/j.hal.2011.10.027.

Ochman, H., Elwyn, S., and Moran, N. A. (1999). Calibrating bacterial evolution. *Proc. Natl. Acad. Sci. U. S. A.* 96, 12638–43. doi:10.1073/pnas.96.22.12638.

Otsuka, S., Suda, S., Li, R., Watanabe, M., Oyaizu, H., Matsumoto, S., et al. (1999). Phylogenetic relationships between toxic and non-toxic strains of the genus *Microcystis* based on 16S to 23S internal transcribed spacer sequence. *FEMS Microbiol. Lett.* 172, 15–21. doi:10.1111/j.1574-6968.1999.tb13443.x.

Schmidt, V. T., Reveillaud, J., Zettler, E., Mincer, T. J., Murphy, L., and Amaral-Zettler, L. A. (2014). Oligotyping reveals community level habitat selection within the genus Vibrio. *Front.*

Microbiol. 5, 1–14. doi:10.3389/fmicb.2014.00563.

Shapiro, B. J., and Polz, M. F. (2014). Ordering microbial diversity into ecologically and genetically cohesive units. *Trends Microbiol.* 22, 235–247. doi:10.1016/j.tim.2014.02.006.