Fall 10-12-2021

Algal Biodiversity Survey of the Middle Chattahoochee River Using DNA Metabarcoding

Amanda Howell

COLUMBUS STATE UNIVERSITY

Algal Biodiversity Survey of the Middle Chattahoochee River Using DNA Metabarcoding

A THESIS SUBMITTED TO

THE COLLEGE OF LETTERS AND SCIENCE

IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

DEPARTMENT OF BIOLOGY

BY

AMANDA L. HOWELL

COLUMBUS, GA

2021

ALGAL BIODIVERSITY SURVEY OF THE MIDDLE CHATTAHOOCHEE RIVER USING DNA METABARCODING

By

Amanda L. Howell

Committee Chair:

Dr. Kevin Burgess

Committee Members:

Dr. Clifton Ruehl

William Kent

Columbus State University

May 2021

Abstract

Aquatic plants play vital roles in water systems by providing ecological services, and one group that is understudied is algae. Algae play a crucial role in water systems as bioindicators and primary producers. They can be challenging to identify with the naked eye, so researchers have been using DNA metabarcoding, where DNA belonging to specific taxa can be isolated and identified from water samples. The objective of this study was to conduct an algal biodiversity survey on the middle Chattahoochee River system in Georgia. Water samples (1L) were filtered and sent to RTL Genomics for sequencing and processing; the data was then analyzed through PAST software to create diversity indices. Overall, time of sampling, lake population, and the interaction between the two showed significance for the following variables: number of individuals, number of taxa, Simpson diversity index, dominance, Shannon diversity index, and evenness. Future recommendations include increasing the number and type of sampling locations and the inclusion of samples collected throughout the year. The preliminary assessment of algae diversity presented here will provide future guidance for water quality management and biodiversity conservation along the middle Chattahoochee River.

Index Words: Biodiversity, DNA barcode, metabarcoding, algae, diversity index, number of individuals, number of taxa, Simpson diversity index, Shannon diversity index, dominance, evenness

Acknowledgments

This study would not have been possible without the help of numerous people and organizations: Columbus Water Works for funding this study; RTL Genomics for processing the algae samples; Kevin Burgess and John Hanson for advising and mentoring; Jennifer Medina, Taylor Bishop, William Kent, Andrew Clark, and John Hall for assisting with water sample collection; Esteban Pinto for helping navigate the diversity data and assistance with the PAST software; and my family, John Hall, Patricia Hall, Jennifer Hood, and Cody Howell, for moral support.

Table of Contents

List of Figures

Figure Page

List of Tables

Table Page

List of Supplemental Figures

Supplemental Figures Page

List of Supplemental Figures

Supplemental Figures Page

List of Supplemental Tables

Introduction

The term biodiversity was first used in 1988 and has become integrated into popular and scientific culture (Titley et al. 2017). Biodiversity can be defined as all variation (genetic, phenotypic, and taxonomic) at all levels of organization (Lovejoy 1997). It can be measured within taxa (genetic diversity; Wilson and Peter 1988), across taxa (species diversity; Solbrig et al. 1994), or even across ecosystems (landscape diversity; Gaston and Spicer 1998). While research has focused on these levels of diversity for land plants (Raymond and Metz 1995, Corlett 2016, Cornwell 2019), animals (Harvey et al. 2006, Balian et al. 2007, McEntee et al. 2020), and insects (García-Robledo et al. 2020, Adams et al. 2020, Crossley et al. 2020), few studies have investigated the diversity of aquatic plants. Aquatic plants (vascular and non-vascular) play vital roles in water systems. They provide habitat for organisms and play an important ecological service through oxygen production (Gettys et al. 2014).

Algae, in particular, are of vital importance for aquatic ecosystems because they play a key role as primary producers in the food chain (Wootton and Power 1993, Shrivastava et al. 2014). In addition, Algae are known to be important bioindicators that can play a critical role in determining the health of aquatic ecosystems (Al-Homaidan et al. 2011, Bellinger and Sigee 2015). Algae, which include blue-green (cyanobacteria), green, red, brown, and diatoms, are generally abundant in aquatic environments like lakes and rivers. Certain types of algae, like blue-green algae, can also cause harm to the water systems forming dense algal blooms, which can contain harmful toxins that are harmful to aquatic life (Falconer et al. 1983, Falconer 1989, Falconer and Humpage 2006). Such algal blooms can also negatively affect water treatment plant operations by hampering filtration and chlorination along with taste and odor concerns in the public's drinking water (McKnight et al. 1983, Beaton and Fine 2018). Algae are either

planktonic (free-floating) or benthic (substrate) within the aquatic environment. However, planktonic algae are more typically found in the main body of standing waters within the epilimnion layer of a stratified water column (Hoek et al. 1995, Wang et al. 2007). Benthic algae are most common in the littoral zone of water bodies. Algal species identification has historically required a light microscope and vast knowledge of algal morphology. However, DNA barcoding technology may prove to be a very effective tool for algal identification. DNA barcoding has developed rapidly over the last decade and has become a valuable tool for surveying biodiversity (Herbert et al. 2003, Kang et al. 2017). This molecular technique is a proven effective tool for fast and accurate species identification (Zou et al. 2016, Chase and Fay 2009). Using molecular tools to identify unidentified algal samples has become a standard technique for species identification in this group of taxa (Le Gall and Saunders 2010, Sherwood et al. 2010, Saunders and McDevit 2012, Hadi et al. 2016, Zou et al. 2016), where morphological identification has several limitations: 1) phenotypic plasticity in traits that are generally used for species recognition can lead to incorrect identifications (Whitman and Agrawal 2009, Belton et al. 2014), 2) morphological methods can overlook cryptic species (Saunders 2008, Radulovici et al. 2010), and 3) a high level of expertise is sometimes required to spot the tiny differences between species (Kim et al. 2014, Manoylov 2014). For example, many green algae species lack noticeable structural features, and the observable traits are very inconstant within the species (FuΩikova et al. 2010). Here, even with a microscope, some species can only be identified based on sexual state (FuΩikova et al.), which may be absent and requires a level of algal expertise that is often lacking.

In addition to traditional DNA barcoding techniques (e.g., those based on Sanger sequencing), recent developments in the field of metabarcoding have been applied to identify algal communities. With metabarcoding, samples can be collected from water systems and then analyzed for different algal species, and with the use of correct primers, can yield data for both cyanobacteria and eukaryotic algae. For example, a study conducted in China (Liu et al. 2020) collected 1-liter surface seawater samples, and after filtration, DNA was extracted and amplified sequenced for the V4 region of the 18S rDNA sequence. Results revealed 326 sequences belonged to eukaryotic algal species; 111 were Ciliophora, 161 were Cercozoa, and 41 were Opisthokonta. Another study conducted in the northern Gulf of Mexico (Bombin et al. 2021) collected water column samples in 1.5 L containers, where 410-450 bp amplicons and 550–590 bp amplicons were generated for the 23S rDNA and part of the large nuclear ribosomal subunit primers (LSU rDNA), respectively. The metabarcoding analyses of both gene regions identified 66 eukaryotic algae, cyanobacteria, and closely related protists species. Overall, such DNA metabarcoding techniques have advantages over simple morphological identification of algae and may be helpful in the identification of algal taxa within freshwater systems.

The overall goal of this study was to use DNA barcoding to conduct an algal biodiversity survey on five lake locations in the middle Chattahoochee river of Georgia. The algal communities currently remain unknown in this section of the river, even though this river meets the surrounding areas' drinking water and recreational needs. Here, I use DNA metabarcoding to identify the diversity (number of individuals, number of taxa, dominance, evenness, Simpsons index, and Shannon diversity index) of algae present in this river system and address the following questions: What are the overall levels of diversity of algal diversity within this river system? 2) Does the algal diversity differ among lakes in the river system? 3) Do measures of

diversity change over time? 4) are the temporal changes in diversity consistent among the lakes (lake by time interaction)? I predict there are only a few species of algae present in the river system and that the algal diversity does not differ among lakes. This prediction is based on the fact that human impact on river water quality has been high, which has ultimately limited the numbers of algal species present, and that all the lakes are within the same river system (i.e., relatively close proximity). Because algae growth is likely to increase across seasons (Round 1984), I also predict that algae diversity will increase over time, but the interaction between location and time will not be significant.

Methods

Study Site:

The Chattahoochee River (hereafter referred to as the Chattahoochee) is a river system within the Apalachicola-Chattahoochee-Flint (ACF) river basin in the southeastern USA that drains an area ~ 19,800 square miles in Georgia, Alabama, and Florida. The Chattahoochee watershed starts above Atlanta in Northern Georgia (Lipford 2004) and flows southwesterly over 400 miles through Georgia and Alabama, where it meets with the Flint River near the Alabama and Florida borders in Lake Seminole (Stephenson 2000). Historically, the Chattahoochee, and the rest of the ACF river basin, have been used for flood control, domestic drinking water, hydroelectric power, transportation, industry, and recreation (Lipford 2004). Numerous dams along the Chattahoochee create reservoirs used for domestic drinking water and power generation. Five of these reservoirs were used for this study. These are located along the middle section of the river and include Walter F. George Lake (E), Goat Rock Lake (GR), Lake Oliver (O), Lake Harding (H), and West Point Lake (WP) (Supplemental Fig. 10).

Sampling:

To assess the diversity of algal communities along the middle section of the Chattahoochee, I selected five sites within each of the five reservoirs as sampling locations (total number of locations sampled = 25). Sites were chosen mainly by accessibility because some of the lakes have fewer public access points (e.g., Lake Oliver and Goat Rock Lake), and all sampling sites were at least 200 feet from each other. The sites were sampled every two weeks in the morning between 8:30 am and 12 pm from May 2018 – October 2018. This period was chosen as the sampling time window because algae tend to move up towards the water's surface

in the morning and sink during the afternoon. The factor of increasing water temperatures may also increase algal production during these months (Round 1984).

At each sampling location, 1L lake water samples were drawn by submerging new, sterile 1-liter polypropylene plastic bottles with a gloved hand below the water surface near the epilimnion layer (Bellinger and Sigee 2015). Each sample was collected near the shore of each sampling site (Deiner and Altermatt 2014), filling each bottle to the neck. The individual samples were stored in an iced cooler while transporting to the lab (which did not exceed 3 hours; Deiner et al. 2015) and then stored in a refrigerator at <2 degrees C until processed for shipment to an offsite Sequencing Facility. The samples were not pooled together but sent out as individual bottles from each site Samples were taken every other week from May – October. Water conditions were at normal flow levels, and the weather conditions were consistent (i.e., no rain within the previous 72 hours before sampling). Samples were then thawed and filtered with Pall Corporation Microfunnel filters. The filters were kept frozen until they were transported to RTL Genomics Laboratory in Houston, Texas (https://rtlgenomics.com/), where they were analyzed using metabarcoding techniques.

Algae Data Analysis and DNA Metabarcoding:

All samples were sequenced for the 23s gene region of the ribosomal genome (see detailed laboratory methods in RTL Genomics 2019; https://rtlgenomics.com/). Each sample was run through the RTL Genomics pipeline to determine the taxonomic information for each read, and then this information was collected for each sample. In general, the data analysis pipeline consists of two major stages: the denoising and chimera detection stage and the microbial diversity analysis stage. During the denoising and chimera detection stage, denoising was

completed using several techniques to remove short sequences, singleton sequences, and noisy reads. After the bad reads were removed, chimera detection was performed to remove the chimeric sequences. Finally, the remaining sequences were then corrected base by base to help remove noise from within the sequence.

Analysis:

Once the quantitative data on the algal composition of each sample was received from the RTL Genomics, data were assessed by lake location, sites within each lake, and date of sampling. The species' count data was then analyzed using the diversity index program PAST 4.03. For this study, I assessed the number of individuals, number of taxa, dominance, evenness, Simpson, and Shannon diversity indices, given that these measures of diversity are often used in other algal studies (e.g., Brown and Bowman 2001, Prasanna and Nayak 2007, Kostryukova et al. 2008). A two-way ANOVA (JMP, Sall et al. 2017) was then used to assess each response variable for variance due to lake location, sampling time and the time X lake location interaction. To conform to the expectations of normality associated with each of the respective ANOVA models, I log-transformed the following response variables, with back-transformed means reported: time/#of individuals, time/dominance, time*population/dominance, population/Simpson, time/Simpson, time*population/Simpson, time/evenness, time*population/evenness.

Results

Overall, 96 species of algae were detected from the samples collected between May-October of 2018 (Table 2). The number of individuals varied among populations and ranged from 2810.37 to 3178.27, with Lake Walter F. George having the highest value (E = 3178.27) and Lake Oliver having the lowest (O = 2810.37). The number of taxa also varied among the populations and ranged from 18.07 to 22.06, with Oliver having the highest value (O = 22.06) and Lake Harding having the lowest (H = 18.07) (Supplemental Table 1). The average number of taxa collected during the nine sampling times range from 11.05 to 30.11 (Table 3), and the average number of individuals collected were 2397.63 to 3315.36 (Table 3). Supplemental Tables 1-9 show an example of a pictorial representation for the average number of taxa (expressed as a proportion) across the nine sampling times, for one of the sampling locations, at one of the lakes sampled.

Results from the 2-way ANOVA showed a number of significant trends for variation in the number of taxa, the number of individuals, dominance, evenness, the Shannon diversity index, and the Simpson diversity index. However, the Simpson diversity index was the only variable that varied significantly among populations and ranged from 0.50 to 0.61, with Goat Rock Lake having the highest values (GR = 0.61) and Lake Harding having the lowest (H = 0.50) (Fig. 1; Table 1).

The time of sampling showed significant variation in the number of individuals, the number of taxa, dominance, Shannon diversity index, Simpson diversity index, and evenness (Table 1). The number of individuals varied significantly among time of sampling and ranged from 2397.63 to 3315.36, with the 9th sampling time being the highest (October) and the 3rd being the lowest (June) (Fig. 2A). The number of taxa varied significantly among time of

sampling and ranged from 11.05 to 30.12, with the 9th sampling time being the highest (October) and the 1st sampling time being the lowest (May) (Fig. 2B). Dominance varied significantly among time of sampling and ranged from 0.34 to 0.58, with the 1st sampling time being the highest (May) and the 5th sampling time being the lowest (July) (Fig. 2C). The Simpson diversity index varied significantly among time of sampling and ranged from 0.42 to 0.66, with the 5th sampling time being the highest (July) and the 1st time of sampling being the lowest (May) (Fig. 2D). The Shannon diversity index varied significantly among time of sampling and ranged from 0.76 to 1.59, with the 5th sampling time being the highest (July) and the 1st sampling time being the lowest (May) (Fig. 2E). Evenness varied significantly among time of sampling and ranged from 0.17 to 0.28, with the 4th sampling time being the highest (June) and the 9th sampling time being the lowest (October) (Fig. 2F).

The interaction between sampling time and population location was significant for the number of taxa, dominance, Shannon diversity index, Simpson diversity index, and evenness (Table 1). The number of taxa seemed to trend higher either towards the hot summer months in the middle of the year or just from the beginning of the growing season to the end (Fig. 3A). Like the number of taxa, dominance seemed to trend higher either towards the summer months in the middle of the year or just from the beginning of the growing season to the end (Fig. 3B). For the Simpson diversity index, E and GR trended to significantly decline at the end of the growing season, while O and H had a significant decline in the middle of the growing season (Fig. 3C). Like the other variables, the Shannon diversity index (Fig. 3D) trended to decline or increase in the middle of the growing season and from the beginning to the end of the year. Evenness also showed a significant decrease from the middle of the summer to the end of the growing season (Fig. 3E).

Discussion

Overall levels of diversity

Overall, I found that algal diversity in the middle Chattahoochee river (~96 taxa) is comparable to other algal studies in similar river systems (e.g., Barinova and Tavassi 2009; Mahadik and Jadhav 2014). In addition, the overall taxonomic composition of the Chattahoochee is also similar to those in the lower Parana river basin of Argentina, where 105 taxa were found (48% of the taxa were Bacillariophyceae, 22% were Cyanobacteria, 18% were Chlorophyta, and the rest (12%) belonged to Euglenophyta, Xanthophyceae, Synurophyceae and Cryptophyta; Rodríguez et al. (2011)). While our numbers are in-line with those found in other studies, there is some evidence that the overall algae diversity of river systems may fluctuate between rainy and dry seasons. For example, Barinova and Tavassi (2009) found 313 algal taxa and cyanobacteria but indicated that taxonomic composition varied among seasons, with diatoms having high numbers in the winter (i.e., wet season) and cyanobacteria and green algae dominating the summer (i.e., dry season). While this discrepancy in numbers between our study and those of the previous study may be due to the timing of our sampling regime (we only sampled in the summer [i.e., dry] season), results collectively underscore that river systems contain a wide variety of algal species, which may buffer the effects of harmful algal blooms due to the proliferation of one or two species. However, this hypothesis remains empirically tested in the Chattahoochee River system.

Diversity across populations

Populations only varied significantly for the Simpson diversity index (SDI), whereas the measures of dominance and evenness did not vary significantly among populations. Collectively

this result is surprising because SDI is a dominance index, meaning it is heavily weighted toward the abundance of the most common species, and it takes into account both richness and evenness (Smith and Grassle 1977, Wilsey et al. 2005, Ma and Ellison 2018). In other words, dominance, evenness, and SDI should show similar trends in our study, where other studies have shown that SDI is indeed correlated with dominance and evenness (Nagendra 2002, Wilsey et al. 2005). One reason for this discrepancy may be that more samples need to be taken from the lakes in my study to show similar trends for dominance, evenness, and SDI. Alternatively, the lack of differences among locations may indicate the true nature of algal diversity along with the river system. Of note, we found that the populations have a similar mean (~0.5) for each of these measures, which is likely influenced by the different lakes being a part of the same river system. The Chattahoochee may have the same diversity spanning the entire river system because each lake population is similarly impacted by anthropogenic disturbance (i.e., urban centers, water recreation). However, these sources of variation remain to be investigated.

Diversity across time

I found that the individuals collected, the number of taxa, dominance, evenness, and the Shannon diversity index varied among sampling times. Overall, the variation in the number of individuals (see Fig. 2) across time did correlate with one other study. Akar et al. (2006) found that the number of individuals had a sharp decline and increased significantly throughout the season. While it is plausible that changes in light intensity and length of day from May to October in our study are likely contributing to the changes in algae production and diversity over time (Castenholz 1960, Bosc et al. 2004), potential environmental sources of variation were not empirically tested in the current study. Our results, however, do highlight that future studies

should not only measure for variance in Algae diversity across the entire growing season (i.e., significant differences were found) but also make every attempt to incorporate any existing data on environmental conditions (e.g., air and water temperature, PH, turbidity) as potential sources of variation across in the context of the metabarcoding results found in this study.

Interactions between location and time for measures of diversity

There was significant variability with the interaction between sampling times and populations for the number of taxa and diversity indices measured in this study. At all the lakes, the beginning of the growing season showed low numbers of taxa, which peaked for all locations around mid-summer. However, values among Lakes began to change in different ways as the season progressed, where some lakes gradually decreased over time, and others fell sharply then increased at the end of the growing season (See fig 3). The interaction between sampling times and populations also significantly varied for dominance, Simpson diversity index, Shannon diversity index, and evenness. For example, although lake diversity values were similar among most of the lakes, Lake Harding, in particular, had a significant decrease in diversity in August and a spike in species dominance (see fig. 3B). While it seems likely that changes in light intensity among lakes may be playing a part in the difference in the number of taxa and diversity at the different lake populations in our study (see also Castenholz 1960, Bosc et al. 2004; Burgmer and Hillebrand 2011), the magnitude of human impact (e.g., recreation) specific to each lake may also be affecting numbers of taxa and algae diversity detected at each of our locations (Lepistö et al. 2004). However, this source of variation remains to be studied. In addition, values may also be decreasing in some months for certain lakes but not others, as specific taxa dominate certain lakes and cause algal blooms that kill off other algal taxa (see also Wehr et al. 2015;

Wurtsbaugh et al. 2019, Amorim et al. 2021). Overall, the interaction between sampling times and populations for the number of taxa and diversity is undoubtedly significant and demonstrates how such values may vary depending on localized environmental conditions (e.g., lake temperature, sunlight, PH, human impact), even though disparate sampling locations (i.e., Lakes) are connected by the same freshwater water system; such sources of variation should be considered in future studies.

Future recommendations

Future recommendations include increasing the number and type of sampling locations and the inclusion of samples collected throughout the year. Moving forward, I recommend more sampling needs to take place in these lake populations. I also recommend weekly sampling to double the number of samples to the data set and sampling in areas with shade and direct sunlight to investigate how diversity varies with light intensity. Current data on algal blooms and environmental sources of variation (e.g., PH, flow, turbidity) must also be tracked to more fully examine interactions between diversity and algal blooms. The preliminary assessment of algae diversity presented here will provide valuable information for the future management of water quality and biodiversity conservation along the middle Chattahoochee River.

References

Adams, B. J., Li, E., Bahlai, C. A., Meineke, E. K., McGlynn, T. P., & Brown, B. V. (2020). Local-and landscape-scale variables shape insect diversity in an urban biodiversity hot spot. *Ecological Applications, 30*(4), e02089.

Akar, B., & Sahin, B. Ü. L. E. N. T. (2006). Benthic algal flora of Karanlik Lake and diversity of epipelic algae.

Al-Homaidan, A. A., Al-Ghanayem, A. A., & Alkhalifa, A. H. (2011). Green algae as bioindicators of heavy metal pollution in Wadi Hanifah Stream, Riyadh, Saudi Arabia. *International Journal of Water Resources and Arid Environments*, *1*(1), 10-15.

Amorim, C. A., & do Nascimento Moura, A. (2021). Ecological impacts of freshwater algal blooms on water quality, plankton biodiversity, structure, and ecosystem functioning. *Science of the Total Environment*, *758*, 143605.

Balian, E. V., Segers, H., Martens, K., & Lévéque, C. (2007). The freshwater animal diversity assessment: an overview of the results. *Freshwater animal diversity assessment*, 627-637.

Barinova, S., & Tavassi, M. (2009). Study of seasonal influences on algal biodiversity in the River Yarqon (central Israel) by bio-indication and canonical correspondence analysis (CCA). *Turkish Journal of Botany*, *33*(5), 353-372.

Beaton, M., and Fine D. (2018). Cyanobacteria and Public Water Systems. Massachusetts Department of Environmental Protection.

Bellinger, E. G., and Sigee, D. C. (2015). Freshwater algae: identification, enumeration and use as bioindicators. John Wiley & Sons.

Belton, G. S., van Reine, W. F. P. H., Huisman, J. M., Draisma, S. G., & D. Gurgel, C. F. (2014). Resolving phenotypic plasticity and species designation in the morphologically

challenging Caulerpa racemosa–peltata complex (Chlorophyta, Caulerpaceae). *Journal of Phycology*, *50*(1), 32-54.

Bombin, S., Wysor, B., & Lopez-Bautista, J. M. (2021). Assessment of littoral algal diversity from the northern Gulf of Mexico using environmental DNA metabarcoding. *Journal of Phycology*, *57*(1), 269-278.

Bosc, E., Bricaud, A., & Antoine, D. (2004). Seasonal and interannual variability in algal biomass and primary production in the Mediterranean Sea, as derived from 4 years of SeaWiFS observations. *Global Biogeochemical Cycles*, *18*(1).

Brown, M. V., & Bowman, J. P. (2001). A molecular phylogenetic survey of sea-ice microbial communities (SIMCO). *FEMS Microbiology Ecology*, *35*(3), 267-275.

Burgmer, T., & Hillebrand, H. (2011). Temperature mean and variance alter phytoplankton biomass and biodiversity in a long-term microcosm experiment. *Oikos*, *120*(6), 922-933.

Castenholz, R. W. (1960). Seasonal Changes in the Attached Algae of Freshwater and Saline Lakes in the Lower Grand Coulee, Washington. *Limnology and Oceanography*, *5*(1), 1-28.

Chase, M. W., and Fay, M. F. (2009). Barcoding of plants and fungi. Science, *325*(5941), 682-683.

Corlett, R. T. (2016). Plant diversity in a changing world: status, trends, and conservation needs. *Plant diversity*, *38*(1), 10-16.

Cornwell, W. K., Pearse, W. D., Dalrymple, R. L., & Zanne, A. E. (2019). What we (don't) know about global plant diversity. *Ecography*, *42*(11), 1819-1831.

Crossley, M. S., Meier, A. R., Baldwin, E. M., Berry, L. L., Crenshaw, L. C., Hartman, G. L., ... & Moran, M. D. (2020). No net insect abundance and diversity declines across US Long Term Ecological Research sites. *Nature Ecology & Evolution*, *4*(10), 1368-1376.

Deiner, K., and Altermatt, F. (2014). Transport distance of invertebrate environmental DNA in a natural river. PloS one, *9*(2), e88786.

Deiner, K., Walser, J. C., Mächler, E., and Altermatt, F. (2015). Choice of capture and extraction methods affect detection of freshwater biodiversity from environmental DNA. Biological Conservation, *183*, 53-63.

Falconer, I. R. (1989). Effects on human health of some toxic cyanobacteria (blue-green algae) in reservoirs, lakes, and rivers. *Toxicity assessment*, *4*(2), 175-184.

Falconer, I. R., & Humpage, A. R. (2006). Cyanobacterial (blue-green algal) toxins in water supplies: Cylindrospermopsins. *Environmental Toxicology: An International Journal*, *21*(4), 299-304.

Falconer, I. R., Runnegar, M. T., & Beresford, A. M. (1983). Evidence of liver damage by toxin from a bloom of the blue-green alga, Microcystis aeruginosa. *Medical Journal of Australia*, *1*(11), 511- 514.

FuΩikova, K., Chien, L., Lewis, L. A., and Karol, K. G. (2010). An assessment of proposed DNA barcodes in freshwater green algae. Cryptogamie, Algologie, *31*(4), 529-555.

Gaston, K. J., and V. I. Spicer. 1998. Biodiversity: an introduction. Blackwell Science, Oxford, UK.

García-Robledo, C., Kuprewicz, E. K., Baer, C. S., Clifton, E., Hernández, G. G., & Wagner, D. L. (2020). The Erwin equation of biodiversity: From little steps to quantum leaps in the discovery of tropical insect diversity. *Biotropica*, *52*(4), 590-597.

Gettys, Lyn A., William T. Haller, and Marc Bellaud. (2014). "Biology and control of aquatic plants." A Best Management Practices Handbook: Third Edition. Aquatic Ecosystem Restoration Foundation, Marietta, GA.

Hadi, S. I., Santana, H., Brunale, P. P., Gomes, T. G., Oliveira, M. D., Matthiensen, A., and Brasil, B. S. (2016). DNA barcoding green microalgae isolated from neotropical inland waters. PloS one, 11(2), e0149284.

Harvey, C. A., Medina, A., Sánchez, D. M., Vílchez, S., Hernández, B., Saenz, J. C., ... & Sinclair, F. L. (2006). Patterns of animal diversity in different forms of tree cover in agricultural landscapes. *Ecological applications*, *16*(5), 1986-1999.

Hebert, P. D., Cywinska, A., Ball, S. L., and Dewaard, J. R. (2003). Biological identifications through DNA barcodes. Proceedings of the Royal Society of London. Series B: Biological Sciences, 270(1512), 313-321.

Heip, C. H., Herman, P. M., & Soetaert, K. (1998). Indices of diversity and evenness. *Oceanis*, *24*(4), 61-88.

Hoek, C., Mann, D., Jahns, H. M., & Jahns, M. (1995). *Algae: an introduction to phycology*. Cambridge university press.

Kang, Y., Deng, Z., Zang, R., and Long, W. (2017). DNA barcoding analysis and phylogenetic relationships of tree species in tropical cloud forests. Scientific reports, 7(1), 12564.

Kim, M. J., Shim, C. K., Kim, Y. K., Hong, S. J., Park, J. H., Han, E. J., ... & Kim, S. C. (2014). Isolation and morphological identification of fresh water green algae from organic farming habitats in Korea. *Korean Journal of Organic Agriculture*, *22*(4), 743-760.

Kononen, K. (2001). Eutrophication, harmful algal blooms and species diversity in phytoplankton communities: examples from the Baltic Sea. *AMBIO: A Journal of the Human Environment*, *30*(4), 184-189.

Kostryukova, A. M., Krupnova, T. G., Mashkova, I. V., Gavrilkina, S. V., and Egorov, N. O. (2018). Phytoplankton diversity in three lakes of South Ural, Russia. *Biodiversitas Journal of Biological Diversity*, *19*(4), 1459-1467.

Lebret, K., Kritzberg, E. S., Figueroa, R., & Rengefors, K. (2012). Genetic diversity within and genetic differentiation between blooms of a microalgal species. *Environmental Microbiology*, *14*(9), 2395-2404.

Le Gall, L., and Saunders, G. W. (2010). Dna barcoding is a powerful tool to uncover algal diversity: A case study of the Phyllophoraceae (Gigartinales, Rhodophyta) in the Canadian flora 1. Journal of phycology, 46(2), 374-389.

Lepistö, L., Holopainen, A. L., & Vuoristo, H. (2004). Type-specific and indicator taxa of phytoplankton as a quality criterion for assessing the ecological status of Finnish boreal lakes. *Limnologica*, *34*(3), 236-248.

Lipford, J. W. (2004). Averting water disputes: A southeastern case study. PERC.

Liu, S., Gibson, K., Cui, Z., Chen, Y., Sun, X., & Chen, N. (2020). Metabarcoding analysis of harmful algal species in Jiaozhou Bay. *Harmful algae*, *92*, 101772.

Lovejoy, T. E. (1997). Biodiversity: what is it. Biodiversity II. Understanding and Protecting Our Biological Resources, 7-14.

Ma, Z., & Ellison, A. M. (2018). A unified concept of dominance applicable at both community and species scales. *Ecosphere*, *9*(11), e02477.

Mahadik, B. B., & Jadhav, M. J. (2014). A preliminary study on algal biodiversity of Ujani reservoir (MS) India. *Bioscience discovery*, *5*(1), 123-125.

Manoylov, K. M. (2014). Taxonomic identification of algae (morphological and molecular): species concepts, methodologies, and their implications for ecological bioassessment. *Journal of phycology*, *50*(3), 409-424.

McEntee, J. P., Burleigh, J. G., & Singhal, S. (2020). Dispersal predicts hybrid zone widths across animal diversity: Implications for species borders under incomplete reproductive isolation. *The American Naturalist*, *196*(1), 9-28.

McKnight, D. M., Chisholm, S. W., & Harleman, D. R. (1983). CuSO 4 treatment of nuisance algal blooms in drinking water reservoirs. *Environmental Management*, *7*(4), 311-320.

Nagendra, H. (2002). Opposite trends in response for the Shannon and Simpson indices of landscape diversity. *Applied geography*, *22*(2), 175-186.

Prasanna, R., & Nayak, S. (2007). Influence of diverse rice soil ecologies on cyanobacterial diversity and abundance. *Wetlands ecology and management*, *15*(2), 127-134.

Radulovici, A. E., Archambault, P., & Dufresne, F. (2010). DNA barcodes for marine biodiversity: moving fast forward?. *Diversity*, *2*(4), 450-472.

Raymond, A., & Metz, C. (1995). Laurussian land-plant diversity during the Silurian and Devonian: mass extinction, sampling bias, or both?. *Paleobiology*, 74-91.

Rodríguez, P., Tell, G., & Pizarro, H. (2011). Epiphytic algal biodiversity in humic shallow lakes from the Lower Paraná River Basin (Argentina). *Wetlands*, *31*(1), 53-63.

Roleček, J., Tichý, L., Zelený, D., & Chytrý, M. (2009). Modified TWINSPAN classification in which the hierarchy respects cluster heterogeneity. *Journal of Vegetation science*, *20*(4), 596-602.

Round, F. E. (1984). *The ecology of algae*. CUP Archive.

RTL Genomics (2019). *Data Analysis Methodology for Microbial Diversity*.

Sall, J., Stephens, M. L., Lehman, A., & Loring, S. (2017). JMP start statistics: a guide to statistics and data analysis using JMP. Sas Institute.

Saunders, G. W. (2008). A DNA barcode examination of the red algal family Dumontiaceae in Canadian waters reveals substantial cryptic species diversity. 1. The foliose Dilsea–Neodilsea complex and Weeksia. *Botany*, *86*(7), 773-789.

Saunders, G. W., and McDevit, D. C. (2012). Methods for DNA barcoding photosynthetic protists emphasizing the macroalgae and diatoms. In DNA barcodes pp. 207-222. Humana Press, Totowa, NJ.

Sherwood, A. R., Sauvage, T., Kurihara, A., Conklin, K. Y., and Presting, G. G. (2010). A comparative analysis of COI, LSU and UPA marker data for the Hawaiian florideophyte Rhodophyta: implications for DNA barcoding of red algae. Cryptogamie Algologie, 31(4), 451.

Shrivastava, A. K., Bharadwaj, M. E. E. N. A. K. S. H. I., and Shrivastava, R. A. N. J. A. N. A. (2014). Algal biodiversity in fresh water reservoir of Durg. Ind J Sci Res, 4(1), 121-126.

Smith, W., & Grassle, J. F. (1977). Sampling properties of a family of diversity measures. *Biometrics*, 283-292.

Solbrig, O. T., H. M. van Emdem, and P. G. W. J. van Oordt. (1994). Biodiversity and global change. CAB International Publishing, Wallingford, UK.

Stephenson, D. S. (2000). The tri-state compact: falling waters and fading opportunities. Journal of Land Use & Environmental Law, 16(1), 83-109.

Titley, M. A., Snaddon, J. L., and Turner, E. C. (2017). Scientific research on animal biodiversity is systematically biased towards vertebrates and temperate regions. PloS one, 12(12), e0189577.

Wang, H., Smith, H. L., Kuang, Y., & Elser, J. J. (2007). Dynamics of stoichiometric bacteria-algae interactions in the epilimnion. *SIAM Journal on Applied Mathematics*, *68*(2), 503-522.

Wehr, J. D., Sheath, R. G., & Kociolek, J. P. (Eds.). (2015). *Freshwater algae of North America: ecology and classification*. Elsevier.

Whitman, D. W., & Agrawal, A. A. (2009). What is phenotypic plasticity and why is it important. *Phenotypic plasticity of insects: Mechanisms and consequences*, 1-63.

Wilsey, B. J., Chalcraft, D. R., Bowles, C. M., & Willig, M. R. (2005). Relationships among indices suggest that richness is an incomplete surrogate for grassland biodiversity. *Ecology*, *86*(5), 1178-1184.

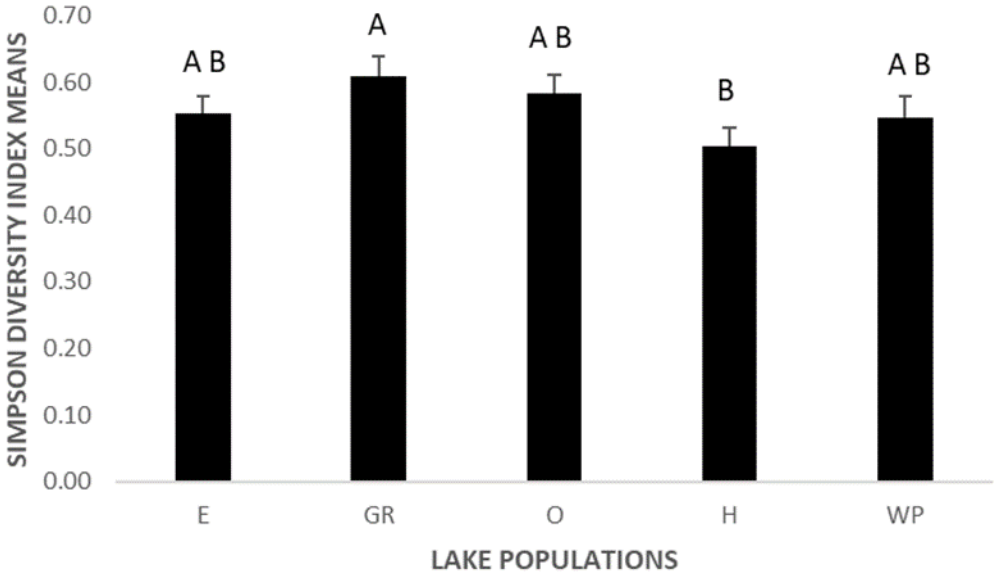
Wilson, E. O., and F. M. Peter. (1988). Biodiversity. National Academy Press, Washington, D.C., USA.

Wootton, J. T., & Power, M. E. (1993). Productivity, consumers, and the structure of a river chain. *Proceedings of the National Academy of Sciences*, *90*(4), 1384-1387.

Wurtsbaugh, W. A., Paerl, H. W., & Dodds, W. K. (2019). Nutrients, eutrophication and harmful algal blooms along the freshwater to marine continuum. *Wiley Interdisciplinary Reviews: Water*, *6*(5), e1373.

Zou, S., Fei, C., Wang, C., Gao, Z., Bao, Y., He, M., and Wang, C. (2016). How DNA barcoding can be more effective in microalgae identification: a case of cryptic diversity revelation in Scenedesmus (Chlorophyceae). Scientific reports, 6, 36822.

Appendix



A

Figure 1. Simpson diversity index of Algae communities collected from each of five lake populations located along the Chattahoochee watershed, West Georgia, USA. Sampling events occurred approximately every two weeks (May – October 2018) during dry weather at lakes Eufaula, Goat Rock, Oliver, Harding, and West Point. Sampling times are as follows: 1 (May); 2-4 (June); 5 (July); 6-7 (August); 8 (September); 9 (October).

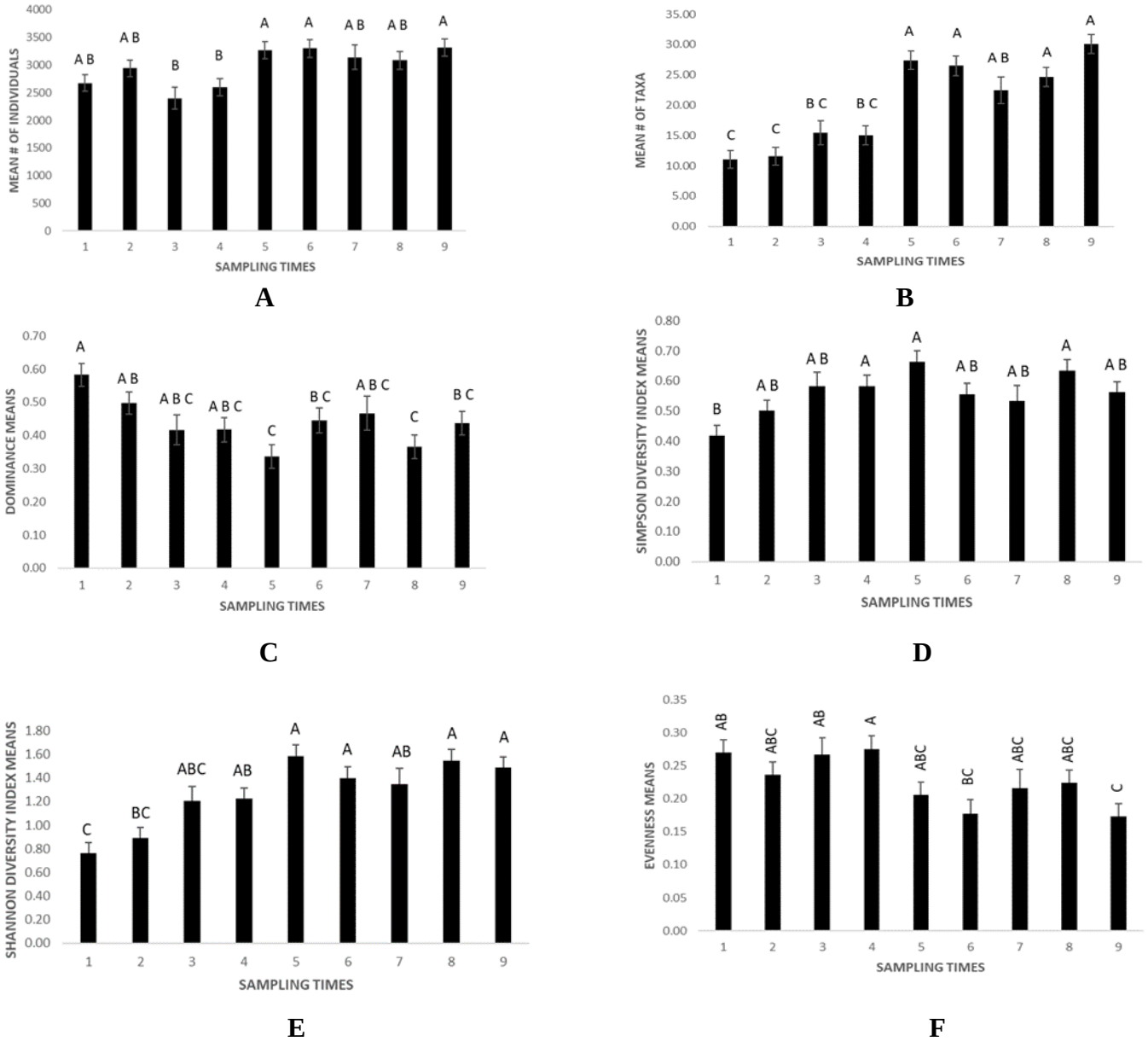
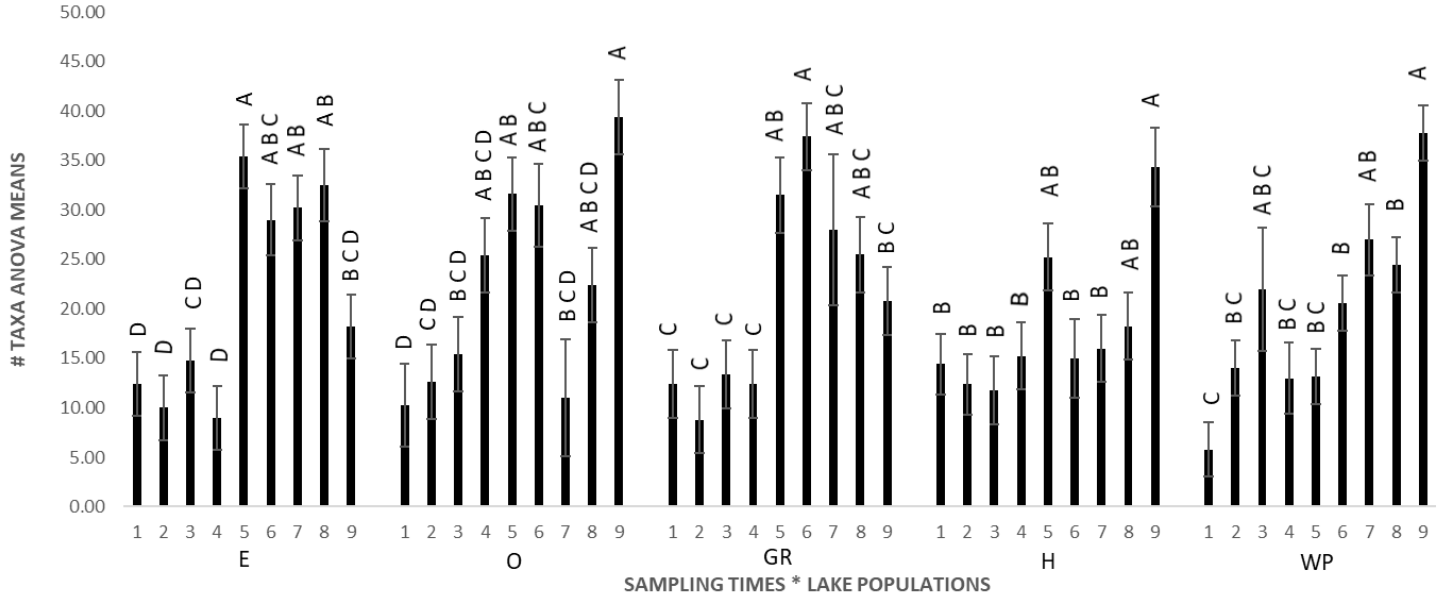
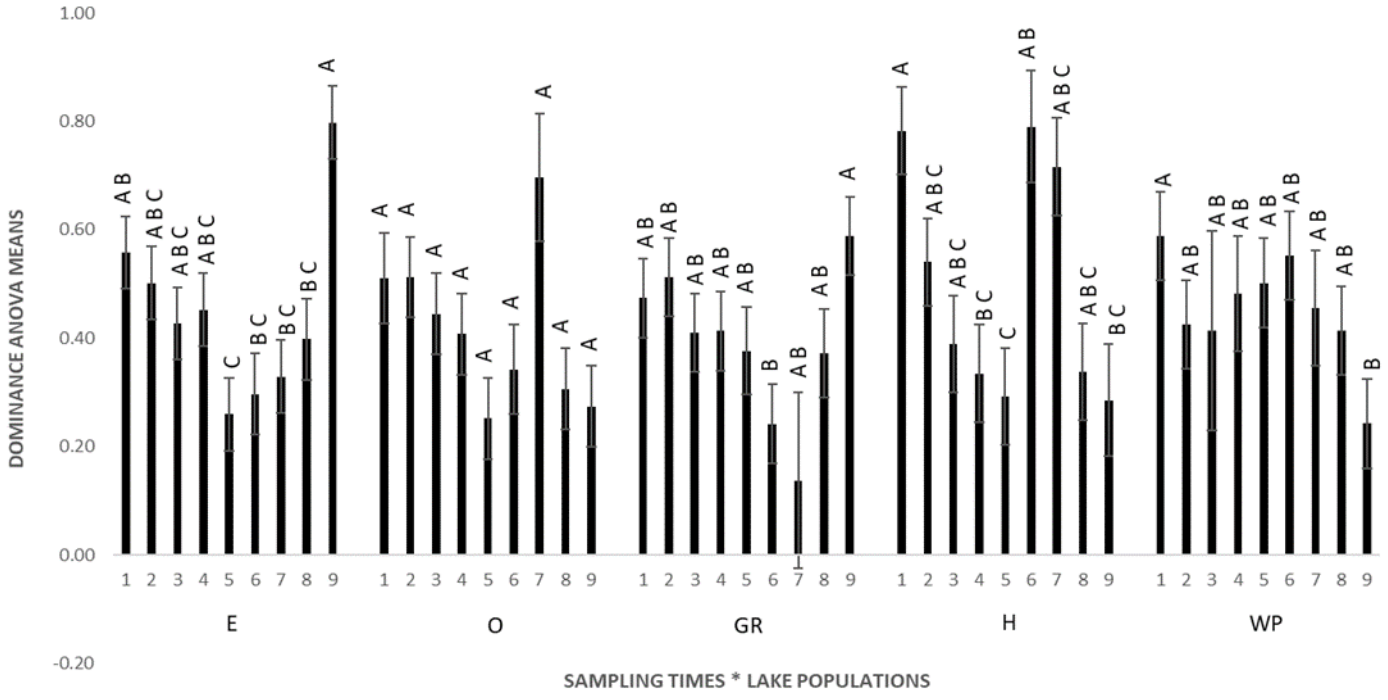


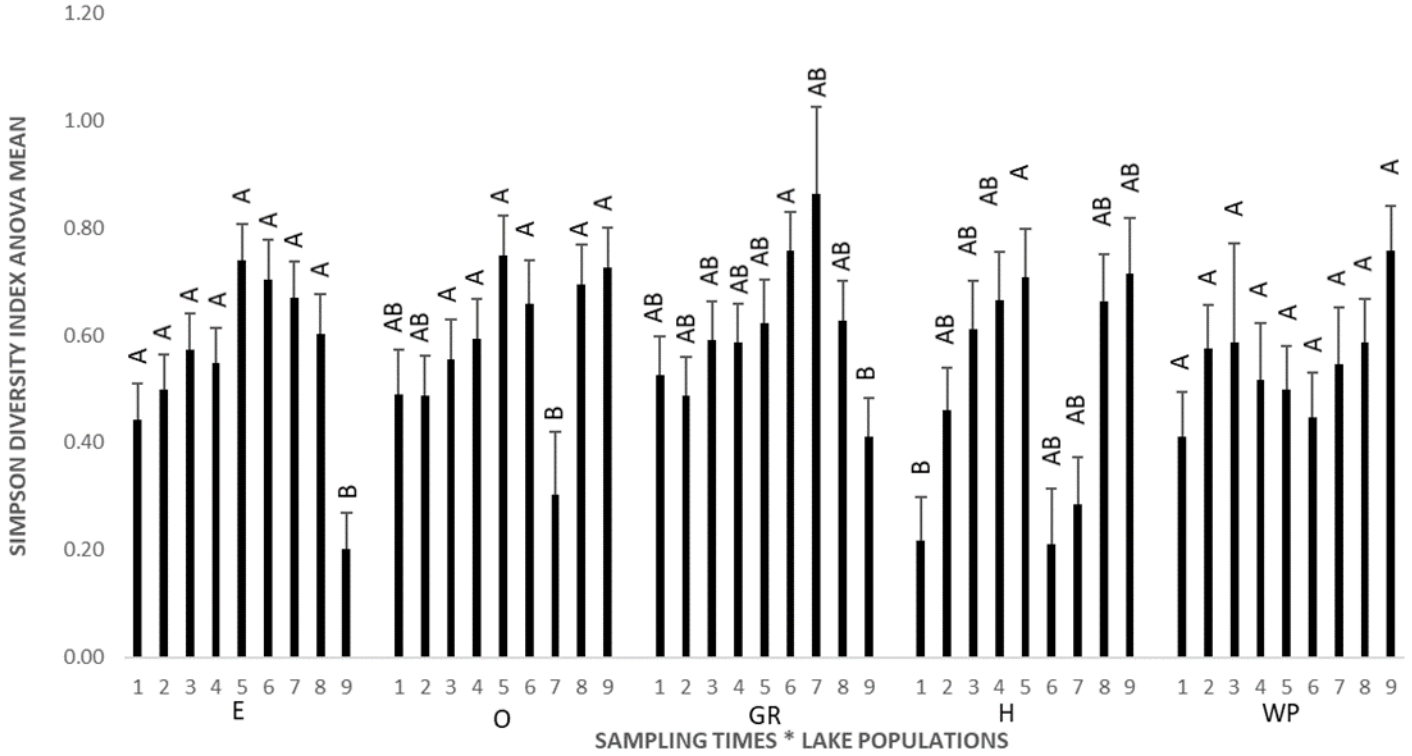
Figure 2: Survey of Algae communities collected from five lake populations located along the Chattahoochee watershed, West Georgia, USA: A) The number of individuals; B) The number of taxa; C) Dominance; D) Simpson diversity index; E) Shannon diversity index; and F) Evenness. Sampling events occurred approximately every two weeks (May – October 2018) during dry weather at lakes Eufaula, Goat Rock, Oliver, Harding, and West Point. Sampling times are as follows: 1 (May); 2-4 (June); 5 (July); 6-7 (August); 8 (September); 9 (October).


TAXA ANOVA MEANS
50.00
45.00
40.00
35.00
30.00
25.00
20.00
15.00
10.00
5.00
0.00
E
O
GR
H
WP
SAMPLING TIMES * LAKE POPULATIONS

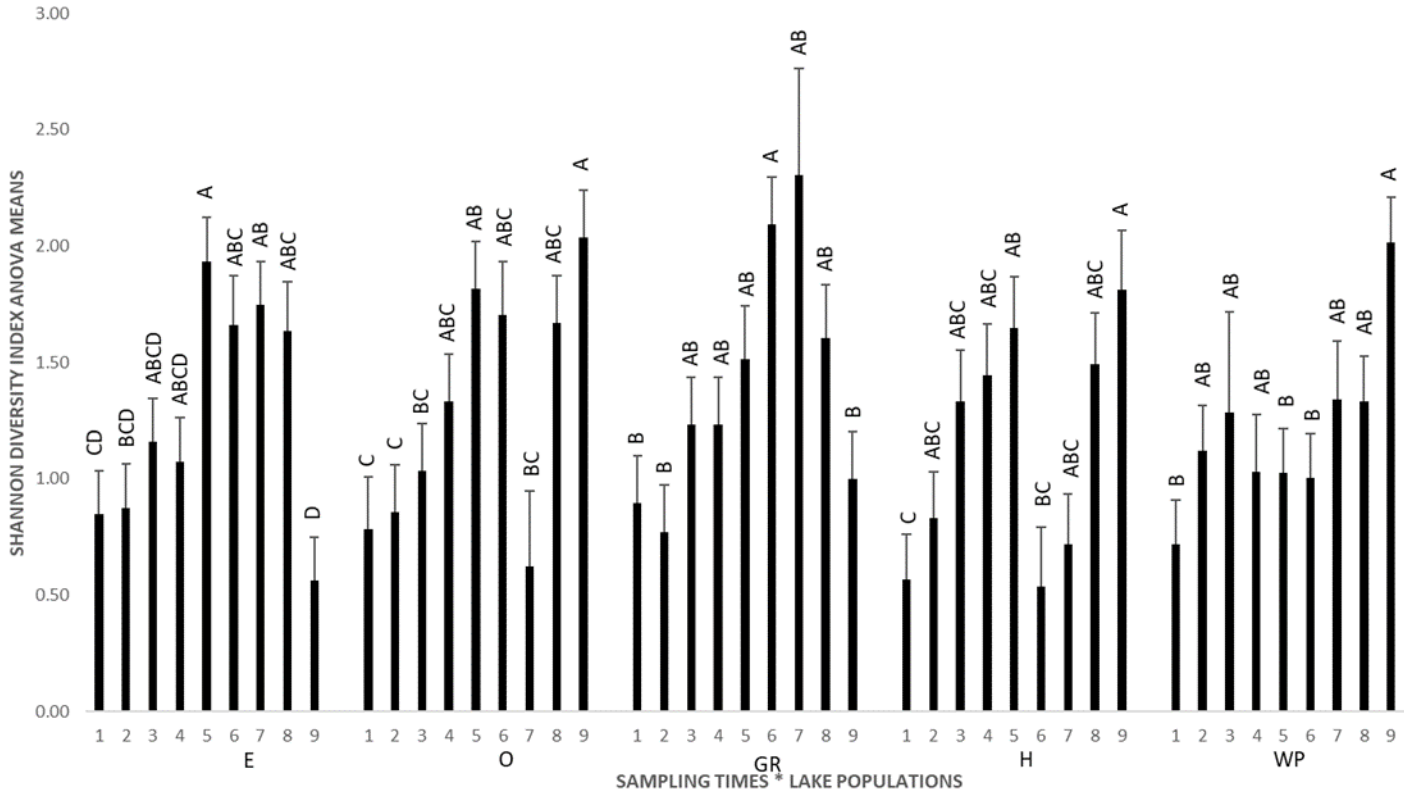
A


DOMINANCE ANOVA MEANS
1.00
0.80
0.60
0.40
0.20
0.00
-0.20
E
O
GR
H
WP
SAMPLING TIMES * LAKE POPULATIONS

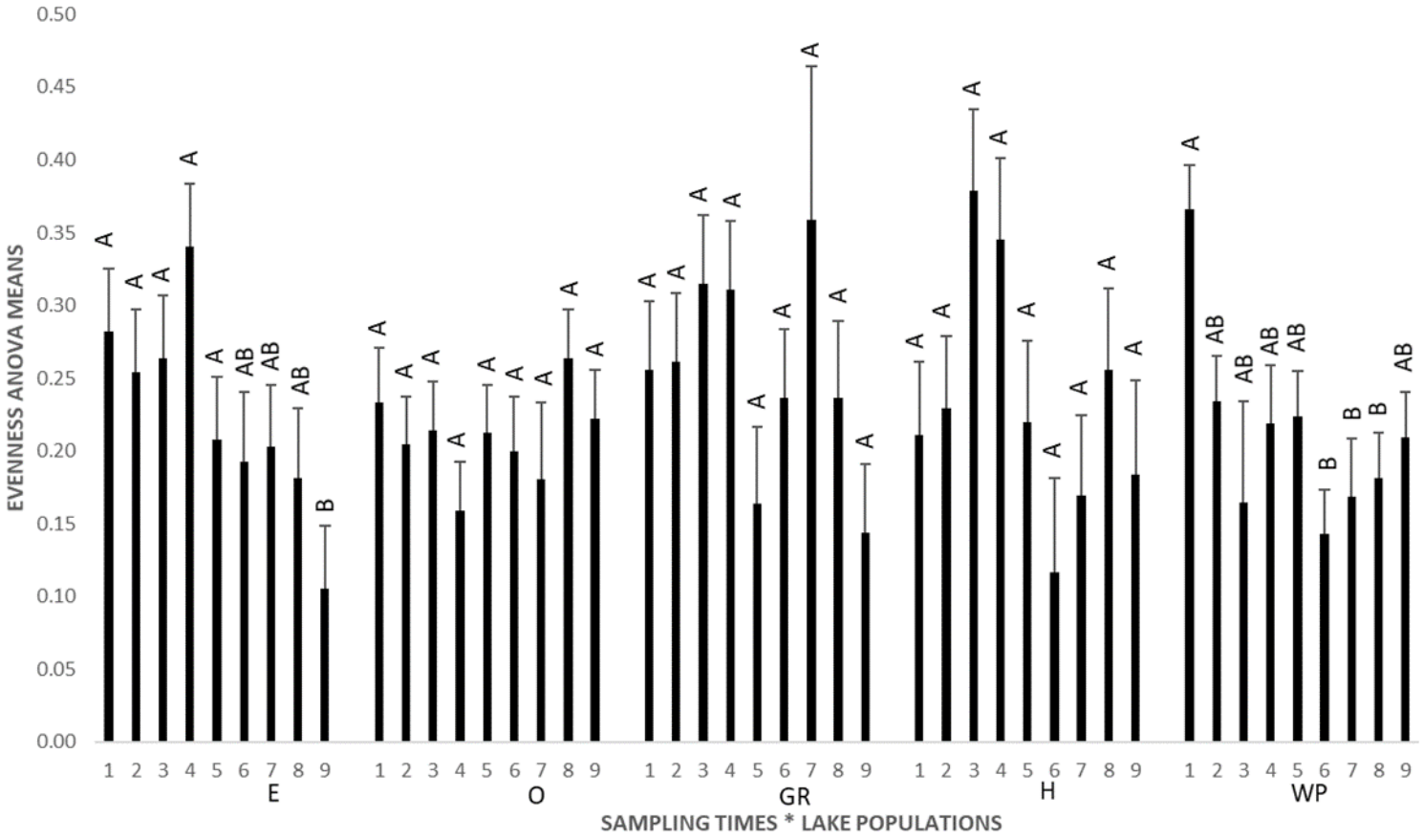
B


SIMPSON DIVERSITY INDEX ANOVA MEAN
1.20
1.00
0.80
0.60
0.40
0.20
0.00
1 2 3 4 5 6 7 8 9
E
O
GR
H
WP
SAMPLING TIMES * LAKE POPULATIONS

C


SHANNON DIVERSITY INDEX ANOVA MEANS
3.00
2.50
2.00
1.50
1.00
0.50
0.00
1 2 3 4 5 6 7 8 9
E
O
GR
H
WP
SAMPLING TIMES * LAKE POPULATIONS

D



E

Figure 3: Difference among five lake populations located along the Chattahoochee watershed, West Georgia, USA. Results of two-way ANOVA for A) the number of taxa, B) dominance, C) Simpson diversity index, D) Shannon diversity index, and E) evenness. Sampling events occurred approximately every two weeks (May – October 2018) during dry weather at lakes Eufaula, Goat Rock, Oliver, Harding, and West Point. Sampling times are as follows: 1 (May); 2-4 (June); 5 (July); 6-7 (August); 8 (September); 9 (October).

Table 1: Results of 2-way ANOVA for number of individuals, number of taxa, dominance, Simpson diversity index, Shannon diversity index, and evenness for Algae populations collected from five lake populations located along the Chattahoochee watershed, West Georgia, USA from May – October 2018.

Effect Tests						
Variable	**Source**	**Nparm**	**DF**	**Sum of Squares**	**F Ratio**	**Prob > F**
# Individuals	Population	4	4	0.05139629	1.2153	0.3067
	time	8	8	0.36881048	4.3603	<0.0001
	time*population	32	32	0.33570606	0.9922	0.4875
# Taxa	Population	4	4	349.5295	1.627	0.1704
	time	8	8	9350.2784	21.7623	<0.0001
	time*population	32	32	6509.7173	3.7878	<0.0001
Dominance	Population	4	4	0.2304445	1.7252	0.1473
	time	8	8	1.3936886	5.2169	<0.0001
	time*population	32	32	3.0934428	2.8949	<0.0001
Simpson	Population	4	4	0.5005327	2.4601	0.0479
	time	8	8	1.3416623	3.2971	0.0017
	time*population	32	32	4.8739474	2.9944	<0.0001
Shannon	Population	4	4	1.21303	1.5684	0.1856
	time	8	8	15.227857	9.8448	<0.0001
	time*population	32	32	22.102885	2.9944	<0.0001
Evenness	Population	4	4	0.0870240	0.7224	0.5779
	time	8	8	0.8825071	3.6628	0.0006
	time*population	32	32	1.7541929	1.8202	0.0090

Table 2: List of algae species detected by DNA metabarcoding in each 1L sample collected along the Chattahoochee watershed, West Georgia, USA from May – October 2018.

Species List
Acetobacter_aceti
Acetobacter_pasteurianus
Achromobacter_arsenitoxydans
Acidisphaera_rubrifaciens
Acidovorax_temperans
Actinomyces_naeslundii
Acutodesmus_obliquus
Aeromicrobium_sp
Agrobacterium_rhizogenes
Agrobacterium_tumefaciens
Anabaena_cylindrica
Ankistrodesmus_stipitatus
Aquabacterium_parvum

Aromatoleum_aromaticum
Asticcacaulis_excentricus
Asticcacaulis_sp
Belnapia_moabensis
Bradyrhizobium_sp
Brevundimonas_naejangsanensis
Brevundimonas_sp
Brevundimonas_subvibrioides
Burkholderia_multivorans
Burkholderia_phenoliruptrix
Burkholderia_sp
Burkholderia_terrae
Burkholderia_ubonensis
Candidatus_Amoebophilus_asiaticus
Caulobacter_segnis
Caulobacter_sp
Caulobacter_vibrioides
Chamaesiphon_minutus
Chlamydomonas_noctigama
Chlamydomonas_peterfii
Chlamydomonas_reinhardtii
Chlamydomonas_sphaeroides
Chlamydomonas_zebra
Chlorella_sp
Chlorella_vulgaris
Chlorellales_unclassified
Chlorogonium_elongatum
Chromulina_sp
Chroococcales_Unclassified
Chroococcidiopsis_sp
Chroomonas_sp
Chrysosporum_ovalisporum
Coccomyxa_subellipsoidea
Coccomyxa_viridis
Coleofasciculus_chthonoplastes
Crinalium_epipsammum
Cryptomonas_curvata
Cryptomonas_marssonii
Cryptomonas_pyrenoidifera
Cryptomonas_sp
Cryptomonas_tetrapyrenoidosa
Cupriavidus_necator
Cyanobium_gracile

Cyanothece_sp
Cylindrospermum_stagnale
Desulfobacterium_vacuolatum
Dinophysis_fortii
Dolichospermum_circinale
Elstera_litoralis
Eukaryota_unclassified
Flavihumibacter_sp
Geitlerinema_sp
Gloeocapsa_sp
Halanaerobiaceae_unclassified
Herbaspirillum_sp
Herminiimonas_arsenicoxydans
Hydrodictyon_reticulatum
Janthinobacterium_lividum
Koliella_longiseta
Lactobacillus_oris
Lactobacillus_paracasei
Lactobacillus_sp
Legionella_pneumophila
Leptolyngbya_fragilis
Leptolyngbya_mycoidea
Leptolyngbya_sp
Magnetospirillum_magneticum
Mallomonas_adamas
Mallomonas_sp
Massilia_niastensis
Massilia_sp
Methanoplanus_Unclassified
Methylobacter_unclassified
Methylomonas_methanica
Methylotenera_mobilis
Methylotenera_versatilis
Methylovorus_sp
Microbacterium_testaceum
Microcystis_aeruginosa
Microthamniales_unclassified
Mobiluncus_Unclassified
Monomastix_minuta
Monoraphidium_sp
Neorhizobium_galegae
Neosynechococcus_sphagnicola
Niastella_koreensis

Oscillatoria_nigro-viridis
Parabacteroides_distasonis
Pedinomonas_sp
Pedobacter_antarcticus
Pedobacter_borealis
Pedobacter_sp
Pelagibacterium_halotolerans
Peridiniopsis_niei
Phormidium_sp
Plagioselmis_nannoplanctica
Pleurocapsa_minor
Polynucleobacter_necessarius
Pseudanabaena_sp
Rhodobacter_sphaeroides
Rhodoferax_ferrireducens
Rickettsia_conorii
Rickettsia_unclassified
Rubrivivax_benzoatilyticus
Scenedesmus_sp
Sinorhizobium_meliloti
Sphaeropleales_unclassified
Sphingobacterium_sp
Spirillum_volutans
Sporohalobacter_unclassified
Stanieria_cyanosphaera
Streptochaeta_angustifolia
Synechococcus_elongatus
Synechococcus_sp
Trebouxia_asymmetrica
Trichormus_azollae
Tropheryma_whipplei
Variovorax_paradoxus
Vaucheria_sp
Weissella_halotolerans
Xanthomonas_campestris
Xanthomonas_translucens
Xanthomonas_unclassified

Table 3: Algae collections across nine sampling times from five lake populations located along the Chattahoochee watershed, West Georgia, USA (May – October 2018). Listed are means for the number of individuals, the number of taxa, dominance, Simpson diversity index, Shannon diversity index, and evenness.

Means			
	Sampling Time	Mean	Standard Error
# of Individuals	1	2671.08	150.10386
	2	2940.8	146.48636
	3	2397.63	199.24299
	4	2596.5633	159.34952
	5	3267.29	153.63619
	6	3301.12	162.68122
	7	3136.7433	221.35119
	8	3084.17	157.08912
	9	3315.36	155.94664
# of Taxa	1	11.05	1.5018937
	2	11.56	1.4656982
	3	15.47	1.9935650
	4	15.01	1.5944030
	5	27.39	1.5372372
	6	26.50	1.6277390
	7	22.44	2.2147729
	8	24.61	1.5717862
	9	30.11	1.5603549

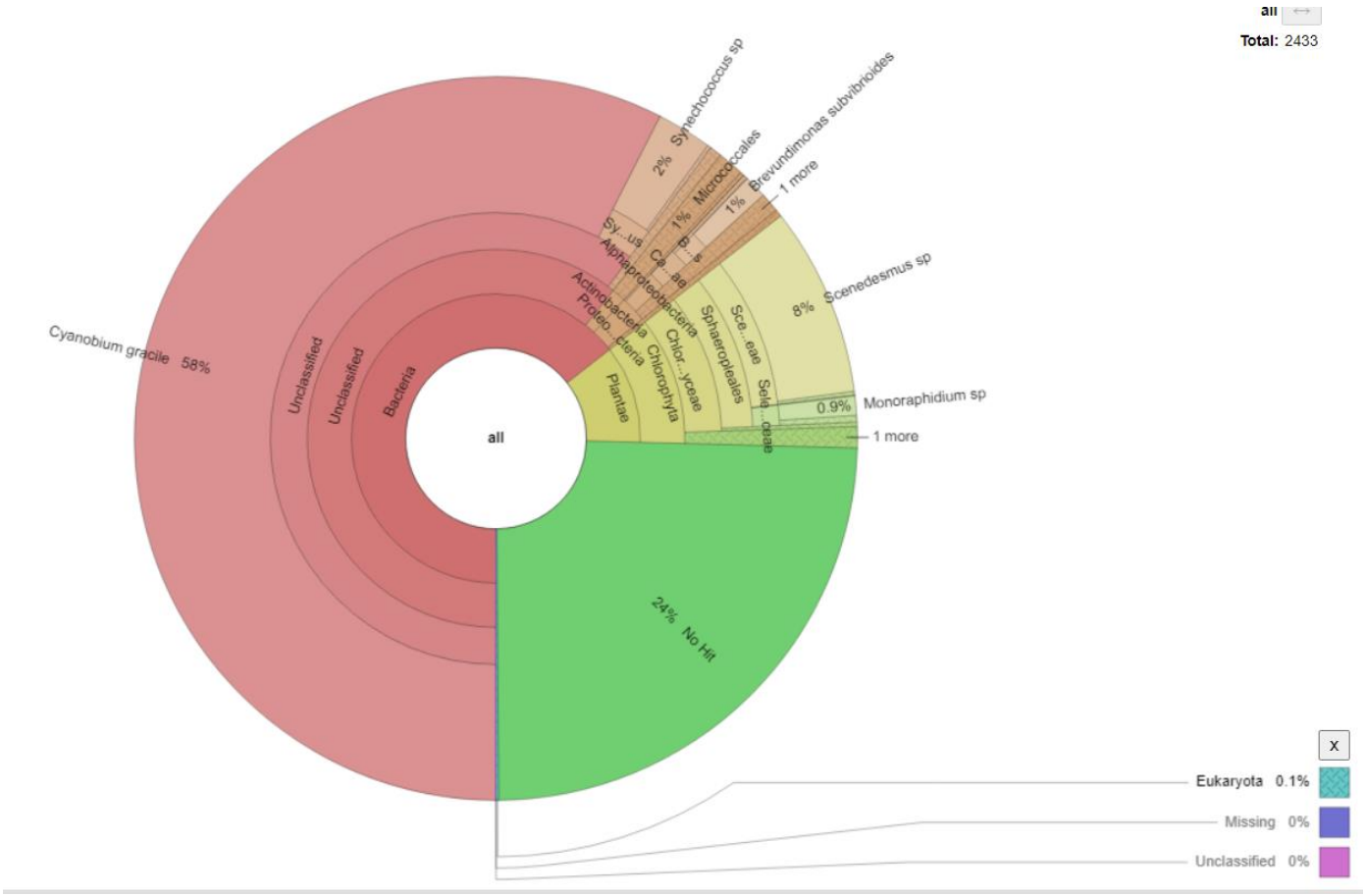


Figure S1: Algae taxa determined by DNA metabarcoding for a 1 Liter water sample collected at Lake Eufaula in the Chattahoochee watershed, West Georgia, USA (Sampling time 1; May).

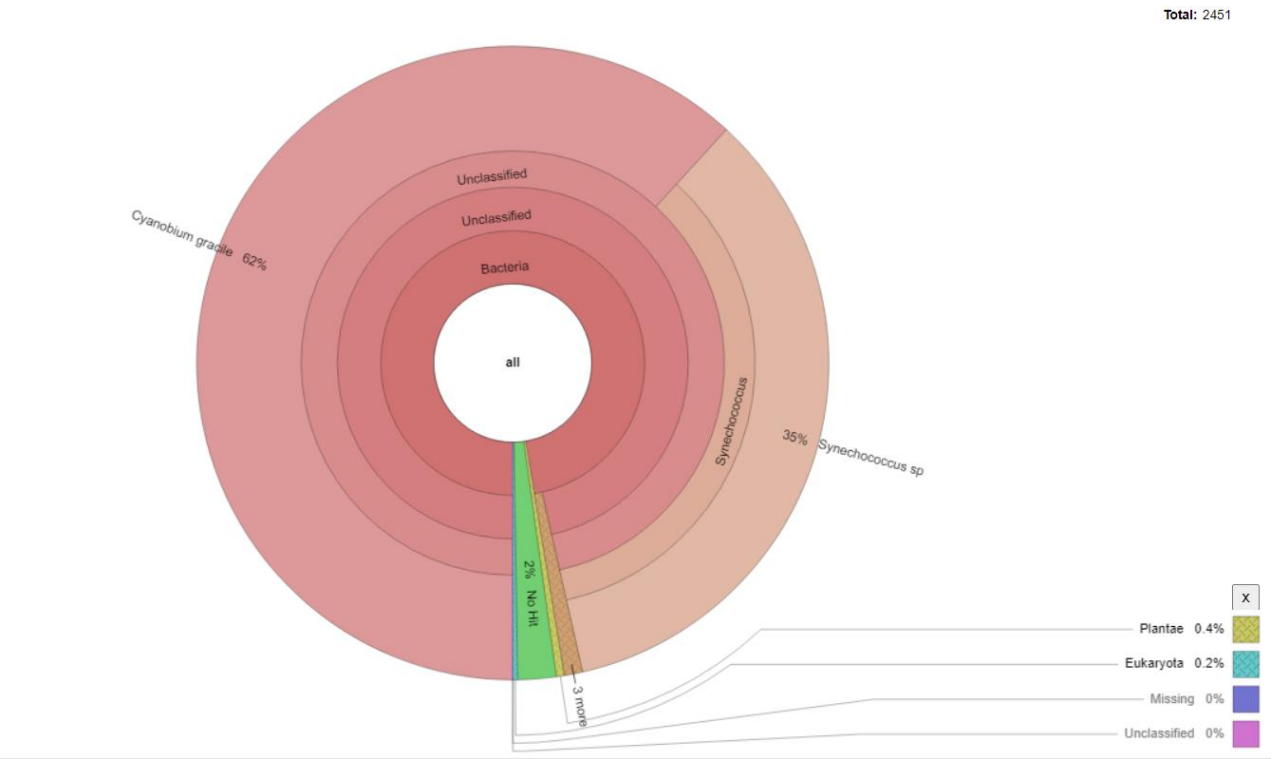


Figure S2: Algae taxa determined by DNA metabarcoding for a 1 Liter water sample collected at Lake Eufaula in the Chattahoochee watershed, West Georgia, USA (Sampling time 2; June).

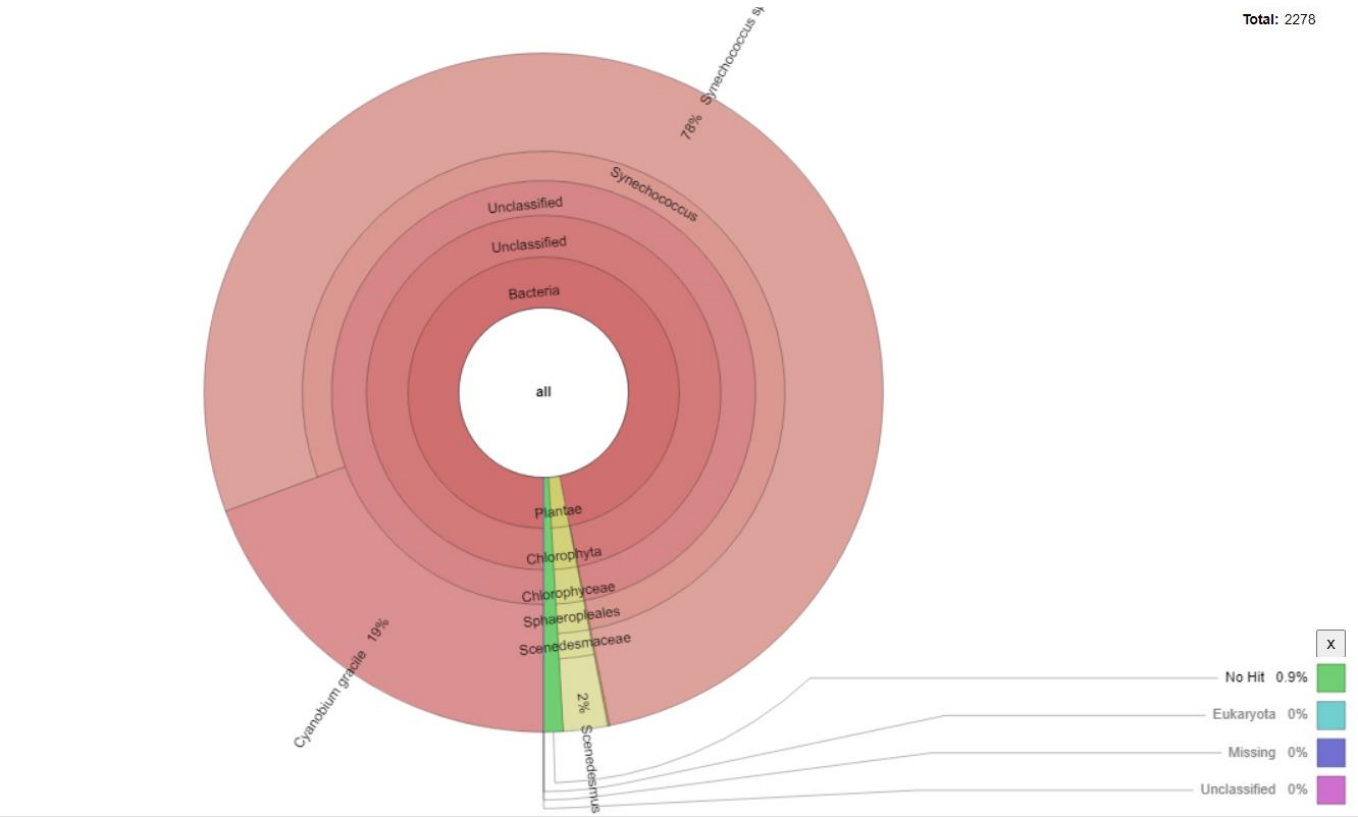


Figure S3: Algae taxa determined by DNA metabarcoding for a 1 Liter water sample collected at Lake Eufaula in the Chattahoochee watershed, West Georgia, USA (Sampling time 3; June).

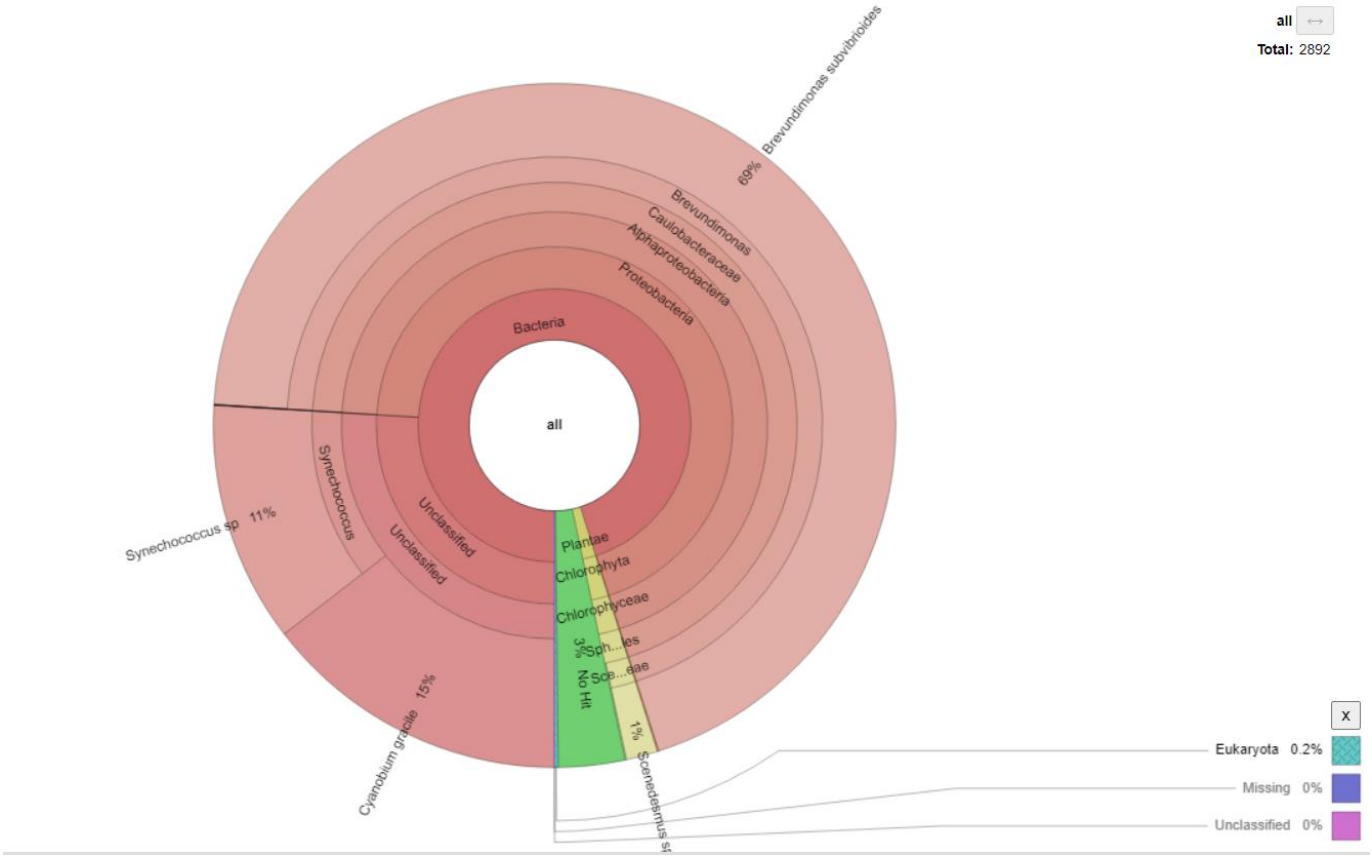


Figure S4: Algae taxa determined by DNA metabarcoding for a 1 Liter water sample collected at Lake Eufaula in the Chattahoochee watershed, West Georgia, USA (Sampling time 4; June).

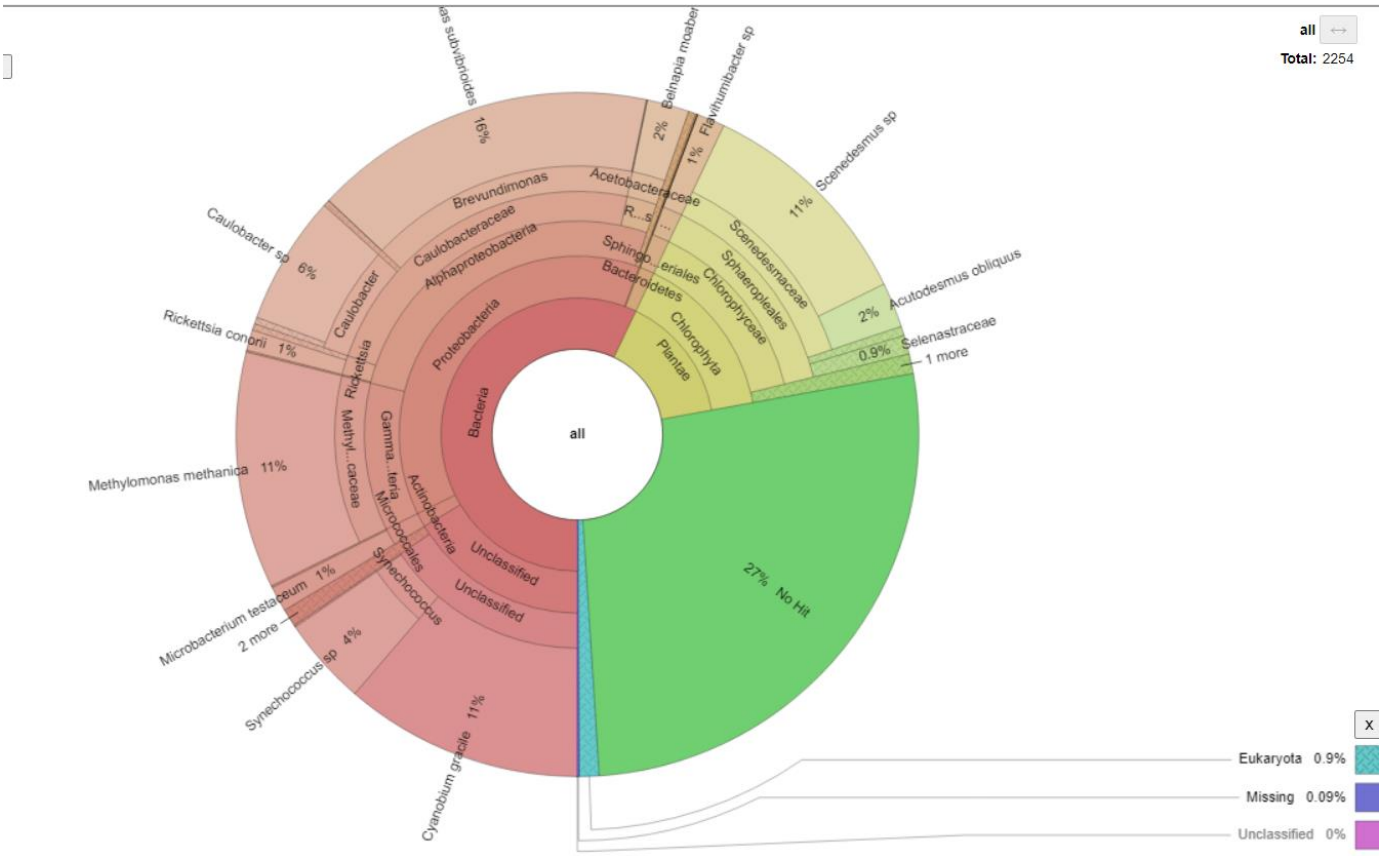


Figure S5: Algae taxa determined by DNA metabarcoding for a 1 Liter water sample collected at Lake Eufaula in the Chattahoochee watershed, West Georgia, USA (Sampling time 5; July).

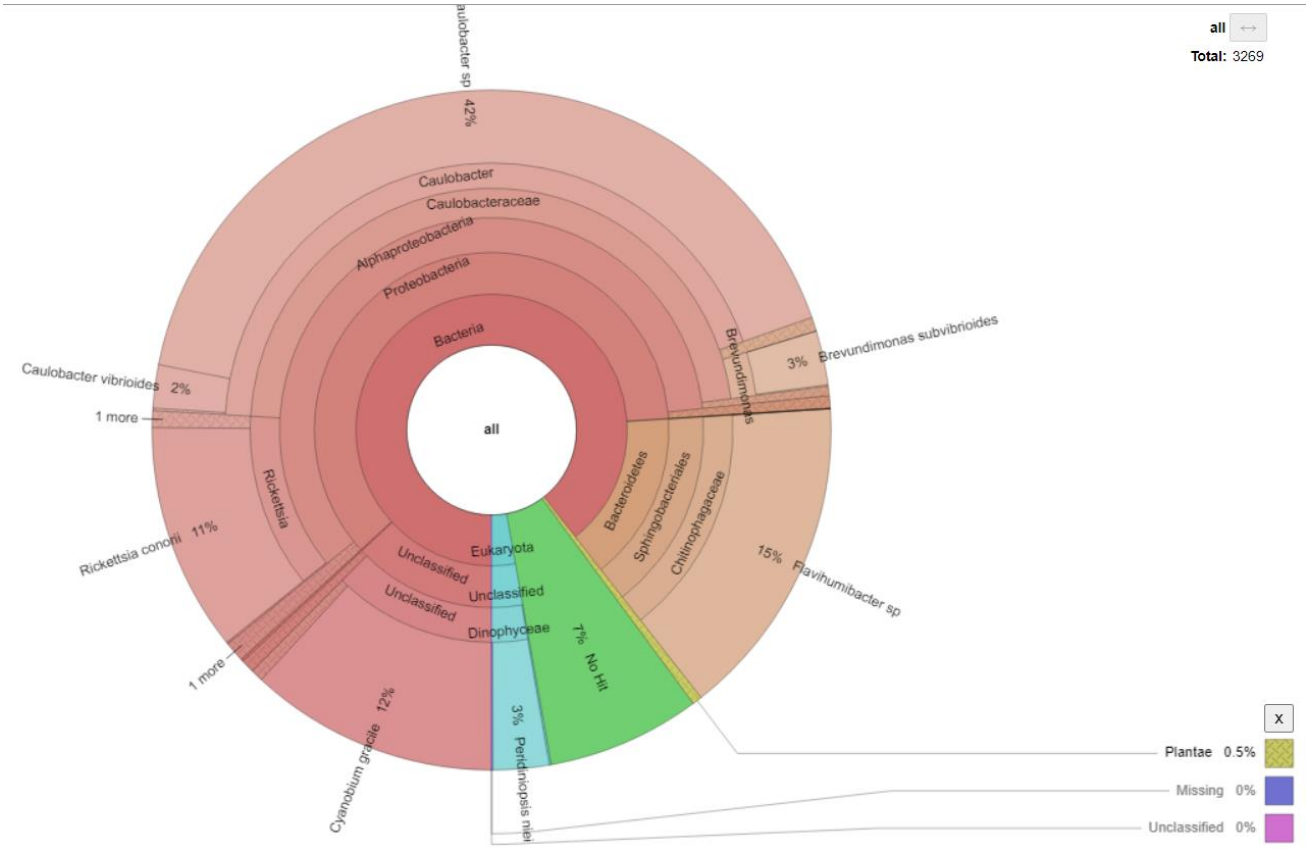


Figure S6: Algae taxa determined by DNA metabarcoding for a 1 Liter water sample collected at Lake Eufaula in the Chattahoochee watershed, West Georgia, USA (Sampling time 6; August).

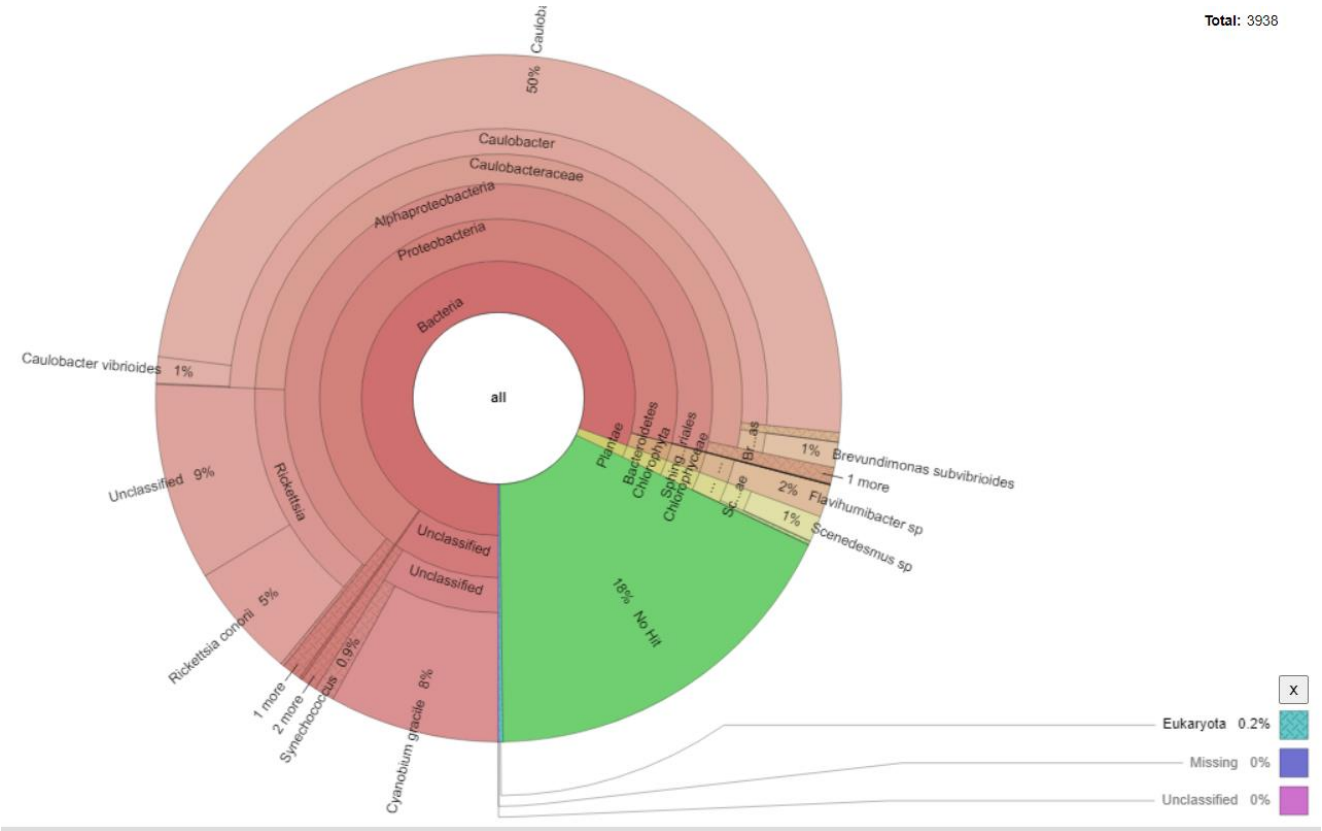


Figure S7: Algae taxa determined by DNA metabarcoding for a 1 Liter water sample collected at Lake Eufaula in the Chattahoochee watershed, West Georgia, USA (Sampling time 7; August).

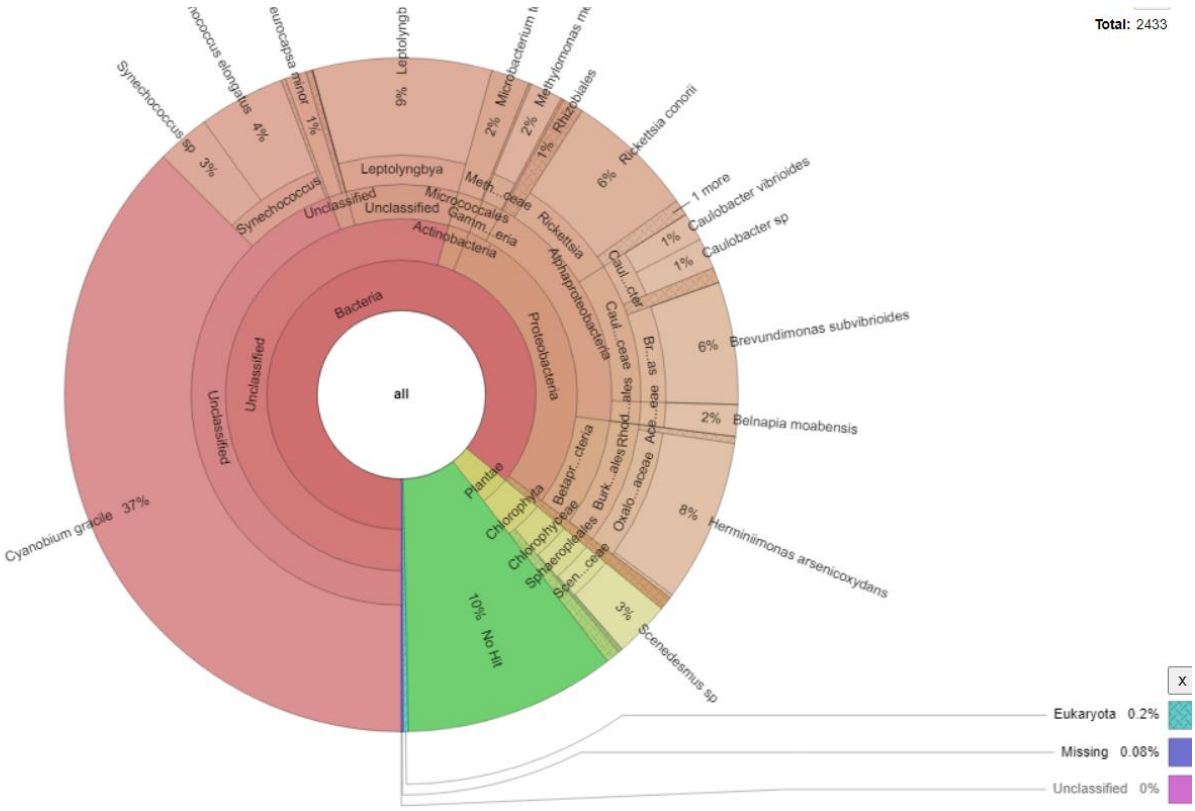


Figure S8: Algae taxa determined by DNA metabarcoding for a 1 Liter water sample collected at Lake Eufaula in the Chattahoochee watershed, West Georgia, USA (Sampling time 8; September).

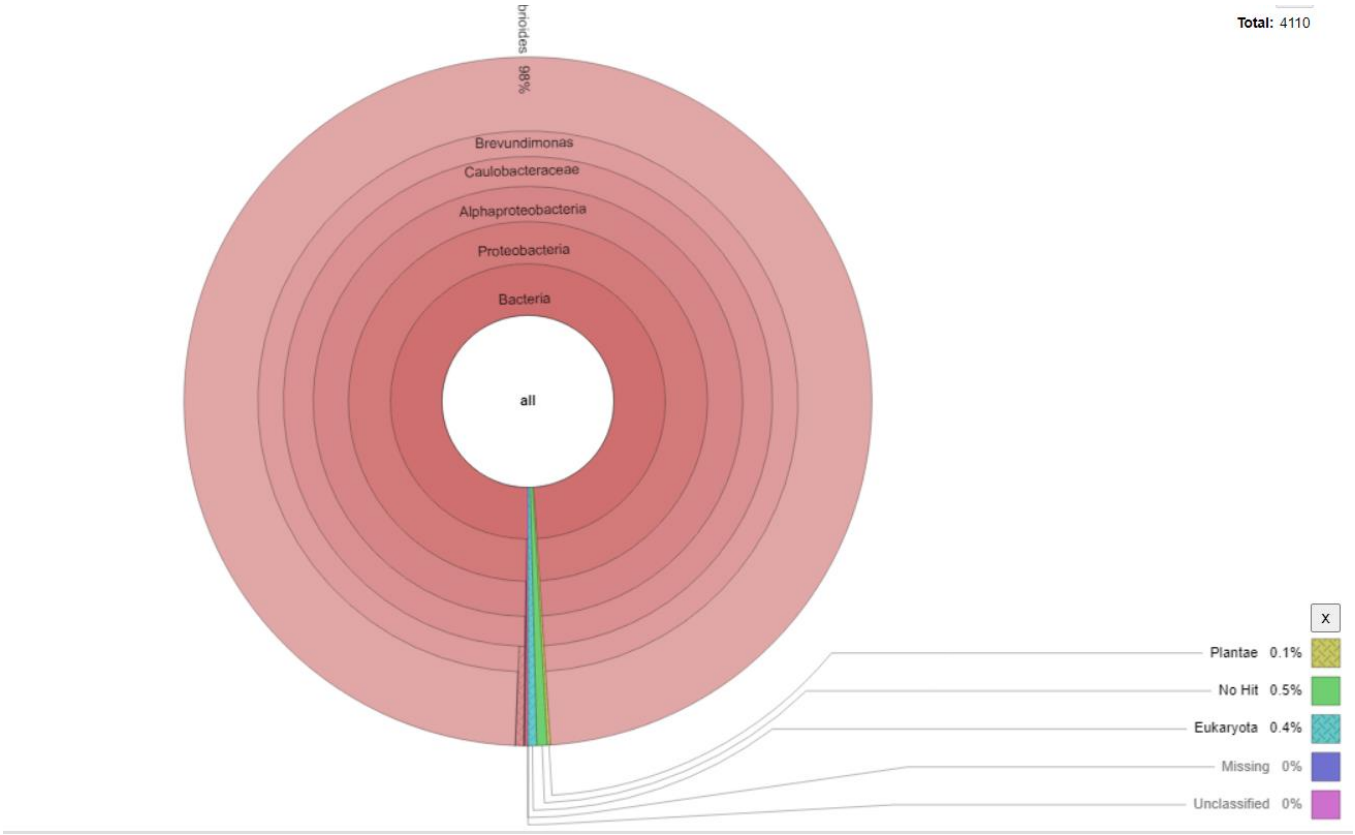


Figure S9: Algae taxa determined by DNA metabarcoding for a 1 Liter water sample collected at Lake Eufaula in the Chattahoochee watershed, West Georgia, USA (Sampling time 9; October).

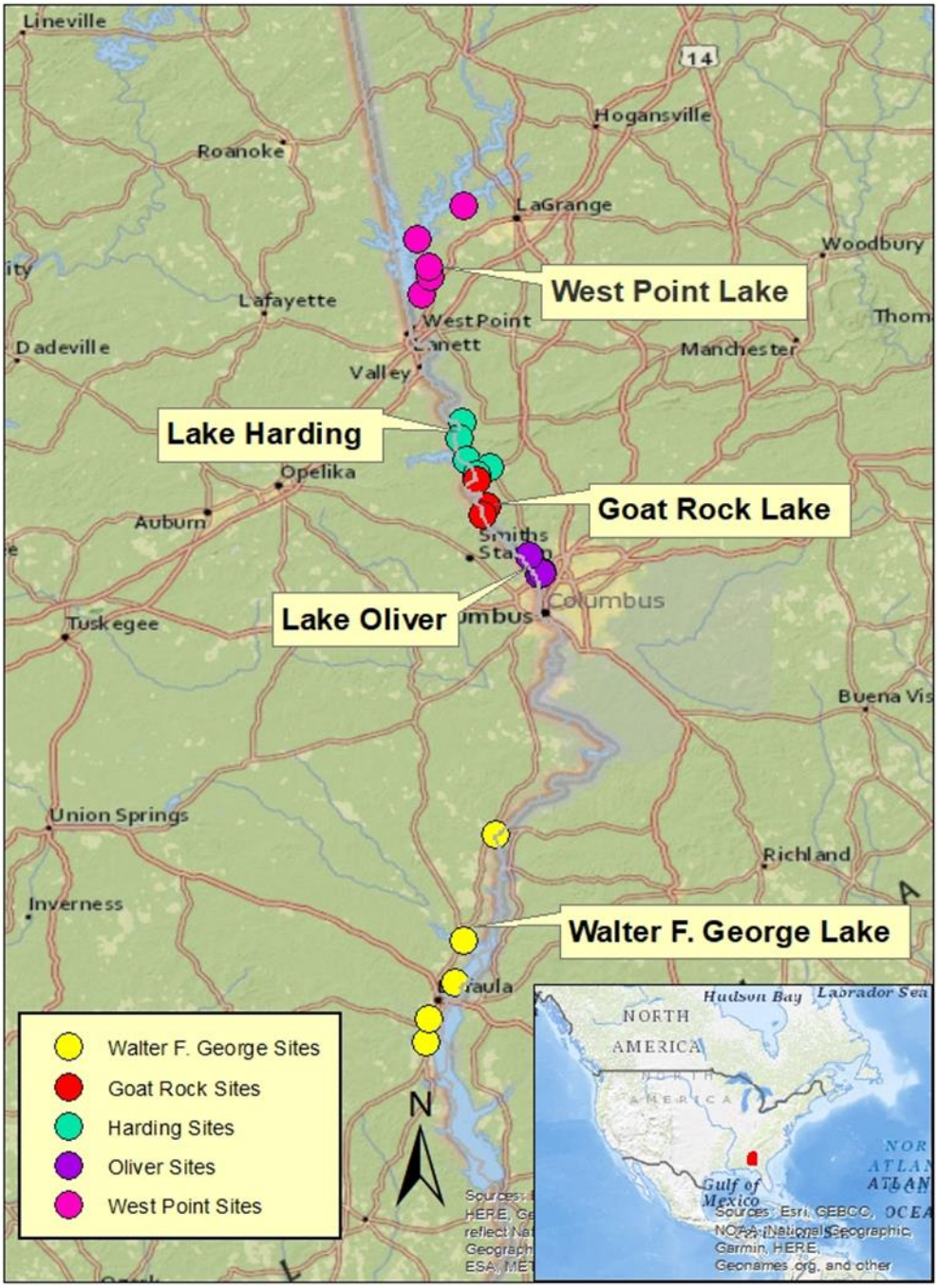


Figure S10: Map of all algae sampling sites in Alabama and Georgia, USA. Five sites for each of the five lakes were selected based on shore-line accessibility in 2018.

Table S1: Algae collections across nine sampling times from five lake populations located along the Chattahoochee watershed, West Georgia, USA (May – October 2018). Depicted are the mean and standard error the for number of individuals, number of taxa, dominance, Simpson diversity index, Shannon diversity index, and evenness.

Variable	Population	Mean	Standard Error	Time	Mean	Standard Error
# Individuals				1	2671.08	150.10386
				2	2940.8	146.48636
				3	2397.63	199.24299
				4	2596.5633	159.34952
				5	3267.29	153.63619
				6	3301.12	162.68122
				7	3136.7433	221.35119
				8	3084.17	157.08912
				9	3315.36	155.94664
# Taxa				1	11.050000	1.5018937
				2	11.560000	1.4656982
				3	15.470000	1.9935650
				4	15.010000	1.5944030
				5	27.390000	1.5372372
				6	26.500000	1.6277390
				7	22.440000	2.2147729
				8	24.610000	1.5717862
				9	30.106667	1.5603549
Dominance				1	0.58244900	0.03441454
				2	0.49843200	0.03358515
				3	0.41689700	0.04568075
				4	0.41792933	0.03653431
				5	0.33617300	0.03522441
				6	0.44469100	0.03729818
				7	0.46683833	0.05074953
				8	0.36555800	0.03601607
				9	0.43759813	0.03575413
Simpson Diversity Index	E	0.55326511	0.02571899	1	0.41755080	0.03441472
	GR	0.60850444	0.03065907	2	0.50156800	0.03358533
	O	0.58360333	0.02767505	3	0.58310300	0.04568099
	H	0.50360600	0.02839938	4	0.58207067	0.03653450
	WP	0.54737296	0.03159117	5	0.66382700	0.03522459
				6	0.55530633	0.03729837
				7	0.53316367	0.05074979
				8	0.63444200	0.03601625
				9	0.56240187	0.03575431
Shannon Diversity Index				1	0.7606296	0.09011468
				2	0.8911080	0.08794292
				3	1.2079790	0.11961531
				4	1.2221417	0.09566530
				5	1.5869930	0.09223532
				6	1.3995413	0.09766549
				7	1.3467317	0.13288793
				8	1.5468430	0.09430828
				9	1.4859253	0.09362239
Evenness				1	0.26950780	0.01904334
				2	0.23654000	0.01858440
				3	0.26706600	0.02527752
				4	0.27492767	0.02021632
				5	0.20548200	0.01949148
				6	0.17765323	0.02063901
				7	0.21571283	0.02808234
				8	0.22374350	0.01992955
				9	0.17291800	0.01978461