Harmful Algae xxx (xxxx) 102757

Contents lists available at ScienceDirect

Harmful Algae

journal homepage: www.elsevier.com/locate/hal

Molecular investigation of harmful cyanobacteria reveals hidden risks and niche partitioning in Kenyan Lakes

Brittany N. Zepernick [a, b, 1], Lauren N. Hart [b, c, d], Emily E. Chase [a], Kaela E. Natwora [b, e], Julia A. Obuya [b, f], Mark Olokotum [b, g], Katelyn A. Houghton [a], E. Anders Kiledal [h], Cody S. Sheik [e], David H. Sherman [d], Gregory J. Dick [h, i, j], Steven W. Wilhelm [a, j], Lewis Sitoki [b, k], Kefa M. Otiso [b, l], R. Michael L. McKay [b, j, m], George S. Bullerjahn [b, j, n, *], Dorine Achieng [b, f], Katelyn B. Barker [b, n], George M. Basweti [b, f], Max Beal [b, o], Katelyn M. Brown [b, n], Aidan Byrne [b, p], Ken G. Drouillard [b, m], Albert Getabu [b, q], Linet I. Kiteresi [b, f], Theodore Lawrence [b, r], Davide Lomeo [b, p], Jared B. Miruka [b, f], Samantha Mohney [b, s], James Njiru [b, f], Pamela Okutoyi [b, k], Reuben Omondi [b, q], Dennis Otieno [b, f], Omondi A. Owino [b, t], Winnie Owoko [b, f], Bethwell Owuor [b, q], Anakalo Shitandi [b, q], Jordyn Stoll [b, u], Miriam Swaleh [b, v], Emma Tebbs [b, p], Emily Varga [b, m], Ryan S. Wagner [b, n], NSF-IRES 2022 Lake Victoria Research Consortium

[a] Department of Microbiology, University of Tennessee Knoxville, Knoxville, TN, USA
[b] NSF-IRES 2022 Lake Victoria Research Consortium, USA
[c] Program in Chemical Biology, University of Michigan, Ann Arbor, MI, USA
[d] Life Sciences Institute, University of Michigan, MI, USA
[e] Large Lakes Observatory, University of Minnesota Duluth, Duluth, MN, USA
[f] Kenya Marine and Fisheries Research Institute, Kisumu, Kenya
[g] National Fisheries Resources Research Institute (NaFIRRI), Jinja, Uganda
[h] Department of Earth and Environmental Sciences, University of Michigan, Ann Arbor, MI, USA
[i] Cooperative Institute for Great Lakes Research (CIGLR), University of Michigan, Ann Arbor, MI, USA
[j] Great Lakes Center for Fresh Waters and Human Health, Department of Biological Sciences, Bowling Green State University, Bowling Green, OH, USA
[k] Technical University of Kenya, Nairobi, Kenya
[l] School of Earth, Environment and Society, Bowling Green State University, Bowling Green, Ohio, USA
[m] Great Lakes Institute for Environmental Research, University of Windsor, Windsor, ON, CA
[n] Department of Biological Sciences, Bowling Green State University, Bowling Green, OH, USA
[o] University of Wisconsin - Madison, Madison, Wisconsin, USA
[p] Department of Geography, King's College, London, United Kingdom
[q] Kisii University, Kisii, Kenya
[r] African Center for Aquatic Research and Education, Ann Arbor, MI, USA
[s] George Mason University, Fairfax, VA, USA
[t] Department of Medical and Applied Sciences, Sigalagala National Polytechnic Kakamega, Kenya
[u] Kent State University, Kent, OH, USA
[v] Technical University of Mombasa, Mombasa, Kenya

ARTICLE INFO

Keywords:
Metabarcoding
Metagenomics
Lake Victoria
Lake Simbi
Lake Naivasha
Methane metabolism
Sulfur limitation

ABSTRACT

Despite the global expansion of cyanobacterial harmful algal blooms (cHABs), research is biased to temperate systems within the global north, such as the Laurentian Great Lakes. This lack of diversity represents a significant gap in the field and jeopardizes the health of those who reside along at-risk watersheds in the global south. The African Great Lake, Lake Victoria, is understudied despite serving as the second largest lake by surface area and demonstrating year-round cHABs. Here, we address this knowledge gap by performing a molecular survey of cHAB communities in three anthropogenically and ecologically important freshwater systems of Victoria's Kenyan watershed: Winam Gulf (Lake Victoria), Lake Simbi and Lake Naivasha. We identified a bloom of non-toxic *Dolichospermum* and toxic *Microcystis* in the Winam Gulf, with data suggesting sulfur limitation shapes com-

* Corresponding author.
E-mail address: bullerj@bgsu.edu (G.S. Bullerjahn).
[1] Current address: Department of Earth, Marine & Environmental Science, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

https://doi.org/10.1016/j.hal.2024.102757
Received 29 August 2024; Received in revised form 4 November 2024; Accepted 9 November 2024
1568-9883/

petition dynamics between these two bloom-formers. Though we did not detect a bloom in Naivasha, it contained the largest diversity of cHAB genera amongst the three lakes. In turn, our results indicated methane metabolism may allow non-toxic picoplankton to outcompete cHAB genera, while suggesting *Synechococcus* spp. serves as a methane source and sink in this system. Lake Simbi exhibited a non-toxic *Limnospira* bloom at the time of sampling with very low abundances of cHAB genera present. Subsequently, these results were employed to design a cHAB screening and risk assessment framework for local stakeholders. Cumulatively, this work serves to increase cHAB research efforts on the international scale while serving as an impetus for cHAB monitoring on the local scale.

1. Introduction

Despite accounting for ~25 % of Earth's unfrozen freshwater, the African Great Lakes remain some of the least-studied to date (Lawrence et al., 2023). Even lesser studied are the thousands of smaller lakes throughout the region which serve as vital resources to human populations and ecological communities alike. A notable threat to these systems are cyanobacterial harmful algal blooms (cHABs) formed by cosmopolitan genera such as *Microcystis*, which are increasing in distribution, duration and frequency due to climate change (Wells et al., 2020; Wilhelm et al., 2020; Zepernick et al., 2023). The present study conducted opportunistic surveys (June-July 2022) to investigate the presence of cHABs within three anthropogenically-impacted and ecologically important Kenyan lakes: Lake Victoria (*African Great Lake*), Lake Simbi (*Soda Crater Lake*) and Lake Naivasha (*Great Rift Valley Lake*).

1.1. Winam Gulf, Lake Victoria (Nyanza Viktoria)

Lake Victoria is the second largest lake in the world by surface area (Hecky, 1993). It is estimated ~50 % of Kenya's population resides within Lake Victoria's basin (Njagi et al., 2022), with the majority of people located within the catchment of Winam Gulf (also referred to as Nyanza Gulf). The Winam Gulf serves as a source of potable water, food and economic revenue to the ~5 million residents within the watershed (McKay et al., 2024; Njuguna et al., 2006; Obuya et al., 2023). Yet, due to deforestation, agricultural intensification and urbanization (Scheren et al., 2000), this freshwater system has become jeopardized by eutrophication (Hecky, 1993; Sitoki et al., 2010). It is difficult to determine when the shift in Lake Victoria's trophic status occurred (Table 1) (Hecky et al., 2010), but symptoms began to manifest in the 1980's. A switch from diatom to cyanobacterial dominance was noted (Gophen et al., 1995), massive "blue-green" algal blooms were attributed to fish kills (Ochumba, 1990) and shifts within the fish fauna were observed (Gophen et al., 1995; Hecky, 1993; Lowe-McConnell, 1994). Notably, heavy blooms of *Microcystis* spp. (Gikuma-Njuru et al., 2006; Krienitz et al., 2002; Ochumba and Kibaara, 1989; Sitoki et al., 2012) and *Dolichospermum* spp. (formerly *Anabaena* spp.) (Brown et al., 2024a; Gikuma-Njuru et al., 2013; Olokotum et al., 2020) have been routinely documented in the Winam Gulf of Lake Victoria. In recent years, studies have confirmed Lake Victoria's cHABs (and their associated toxins) pose a significant risk to human health and confer negative socioeconomic consequences (Brown et al., 2024a; Obuya et al., 2023; Olokotum et al., 2020; Roegner et al., 2020; Roegner et al., 2023). Yet, to date, cHABs within Lake Victoria have not been widely characterized at a molecular level (Lawrence et al., 2023; Olokotum et al., 2020).

Table 1
Characteristics of the three lakes that were opportunistically sampled during the June-July 2022 *NSF IRES Advanced Studies Institute on Water Quality and Harmful Algal Blooms in Lake Victoria, Kenya*. Asterisks indicate values reported correspond to those recorded upon sample collection in this study. Otherwise, values are the most recently reported in the literature (Ballot et al., 2005; Gharib et al., 2023; Ojiambo and Lyons, 2019; Opiyo et al., 2019; Romero and Alexander, 2006; Sitoki et al., 2012; Yongo et al., 2023).

Characteristics	Lake Victoria (Winam Gulf)	Lake Simbi	Lake Naivasha (Crescent Island)
Location of sample collection*	0°31'17.63" S 34°27'25.30" E	0°22'10.81" S 34°37'44.11" E	0°45'31.00" S 36°25'30.00" E
Surface Area (km^2)	1,300	0.29	145
Mean Depth (m)	10	23	5
Residence time (yr)	1-1.5	Unknown	3-6
Altitude above sea level (m)	1,134	1,142	1,890
Population in basin	5 million	250,000	380,000
Conductivity ($\mu S \bullet cm^{-1}$)	100-200	16,000	240-400
pH*	7.4	10.34	7.4
Trophic State Indices	HE, TSI = 145	HE, TSI = 87.01	E/HE, TSI = 66
Potable water source?	Yes	No	Yes
cHAB visibly present?*	Yes, Algae Torch	Yes, microscopy	No

1.2. Lake Simbi (Simbi Nyaima)

Lake Simbi is a soda crater lake located ~1 km from the Winam Gulf of Lake Victoria. The lake's physiochemistry is consistent with a typical soda lake: eutrophic, stratified, hypoxic and routinely surpassing pHs of 10 (Table 1) (Ballot et al., 2005; Opiyo, 2020; Opiyo et al., 2019). The planktonic community of Simbi is dominated by cyanobacterial blooms of *Limnospira* spp. (formerly *Arthrospira spp.*) (Ballot et al., 2005), with historical studies describing Simbi as a "unialgal cyanobacterial bloom of *Limnospira* spp." (Finlay et al., 1987; Melack, 1979). While the alkaline-saline chemistry of the lake renders it unsuitable for direct human use, Simbi is a vital cultural, economic and ecological reserve (Hayombe et al., 2014). Notably, Simbi hosts populations of the endangered Lesser Flamingo (*Phoeniconaias minor*) (Tuite, 1979) which feeds on *Limnospira* as its primary food source (Raini, 2006). As a result, Simbi generates considerable revenue *via* environmental tourism (Hayombe et al., 2014). However, agricultural development in recent decades has exacerbated eutrophication (Opiyo et al., 2019). Coincidently, recent studies have reported alarming declines in flamingo populations (Ndetei and Muhandiki, 2005; Oduor, 2018) and marked shifts in the cyanobacterial community. Notably, *Microcystis* spp. dominated Simbi's water column throughout a six-month period in 2018-2019 (Opiyo, 2020) and Ballot et al. (2005) detected substantial levels of microcystins (19.7-39.0 $\mu g \bullet g^{-1}$ dry weight of microcystin-LR) in Simbi throughout June 2001-2002. In turn, Ballot et al. (2005) raised concerns that *Limnospira* species within Simbi may possess the ability to synthesize microcystins. Yet, despite substantial shifts in cyanotoxin concentrations and community composition, there have been no further investigations, to our knowledge, regarding cHABs within this lake.

1.3. Lake Naivasha

Naivasha is the second largest freshwater lake in Kenya (Renaut and Owen, 2023), and is one of seven lakes located within the Great Rift Valley. The physiochemistry of Naivasha is distinct from its neighbors. In contrast to the alkaline-saline soda lakes within the Great Rift Valley, Naivasha is characterized by nutrient-poor waters, low-salinities (*i.e.,* 282-374 $\mu S \bullet cm^{-1}$) and slightly basic pHs (~8-9) across both wet and dry seasons (Ballot et al., 2009; Harper et al., 2011; Otiang'a-Owiti and Oswe, 2007; Yongo et al., 2023). As a result, the fresh waters of Naivasha in a region otherwise dominated by soda lakes elevates the importance of this system to surrounding communities. Naivasha pro-

vides drinking water, food, and irrigation to ~380,000 residents within its watershed (Onywere et al., 2012; Yongo et al., 2023). Yet, rapid development has threatened this critical resource in recent decades. Notably, the region is one of the fastest growing in Kenya, experiencing a population increase of 137 % from 1999-2009 (Onywere et al., 2012). Concurrently, Naivasha has undergone rapid physiochemical and biotic changes in recent decades. For example, the surrounding watershed has undergone intense agricultural development as foreign large-scale farming operations replace local small-scale establishments (Harper et al., 2011). Cumulatively, the lake has experienced progressive eutrophication (Harper et al., 2011; Hubble and Harper, 2002; Otiang'a-Owiti and Oswe, 2007) - shifting its status to eutrophic (Table 1) (Harper et al., 1993; Yongo et al., 2023). Coinciding with shifts in trophic status – a shift in phytoplankton composition has occurred. Surveys conducted throughout 2001-2005 noted cyanobacterial dominance and a dense bloom of *Microcystis aeruginosa* was reported in 2006 (Ballot et al., 2009). Most recently, a study conducted from February-July 2019 found cyanobacteria continued to dominate the water column (Owino et al., 2020). However, to our knowledge, there have not been any further studies concerning cHABs within Naivasha.

1.4. Investigating the presence of cHABs and genomic potential for toxicity

Though each possesses a distinct physiochemical profile (Table 1), Winam Gulf (Lake Victoria), Lake Simbi and Lake Naivasha are unified by acute eutrophication and episodic reports of cHABs. Despite this risk, the knowledge that is presently available is largely based on microscopy – which falls short of determining if the cyanobacteria possess genes associated with toxin biosynthesis. More broadly, there has been a lack of molecular characterization concerning these lakes to date, representing a considerable gap within the field. In this study, we utilized genomic and bioinformatic techniques to 1) Establish baseline molecular characterizations of the cyanobacterial community, 2) Determine the toxic biosynthetic potential of cyanobacteria and 3) provide baseline resources and recommendations for future risk assessment in these vulnerable freshwater systems.

2. Methods

2.1. Sample collection and water column physiochemistry

Samples of opportunity (n = 26) were collected from the planktonic water column across Lakes Victoria (n = 22), Simbi (n = 2) and Naivasha (n = 2) in June-July of 2022 as part of the *NSF IRES Advanced Studies Institute on Water Quality and Harmful Algal Blooms in Lake Victoria, Kenya (2022-2023)* (Zepernick et al., 2024 in review). To investigate the presence of cHABs in these lakes, six libraries were selected for analysis originating from Homa Bay (Winam Gulf), (0° 31'17.63" S, 34° 27'25.30" E), Lake Simbi (0° 22'10.81" S, 34° 37'44.11" E) and Crescent Island, Lake Naivasha (0° 45'31.00" S, 36° 25'30.0" E). Sample sites spanned various trophic states and bloom severity (Table 1) (Fig. 1) (Supplemental Table 1).

Whole water samples (1 L) were collected in biological duplicate from a depth of ~0.5 m. Subsequently, samples for DNA analysis were collected onto sterile 0.22 μm pore-size Sterivex filters (Sigma Aldrich, St. Louis, MO, USA) and stored at room temperature in DNA/RNA Shield (Zymo Research, Irvine, CA, USA) until extraction. Filtrate was collected in biological duplicate for dissolved nutrient analysis (NO_X, NH_4, DRP, SiO_2) and stored at -20 °C until processing on a QuAAtro 5-Channel continuous segmented flow auto-analyzer (Seal Analytical Inc., Mequon, WI, USA). A multiparameter sonde (YSI Inc., Yellow Springs, OH, USA) and Algae Torch (bbe Moldaenke, Kiel, Germany) were available for *in situ* measurements at Lake Victoria. Otherwise, the temperature and pH of each system was measured *in situ* using an LCD digital aquarium thermometer (Vivosun, Ontario, CA, USA) and high accuracy MQuant pH strips (0.3 intervals, range pH 6.5-10) (Millipore Sigma, Burlington, MA, USA). All metadata are available online at the Biological and Chemical Oceanography Data Management Office (Bullerjahn et al., 2024).

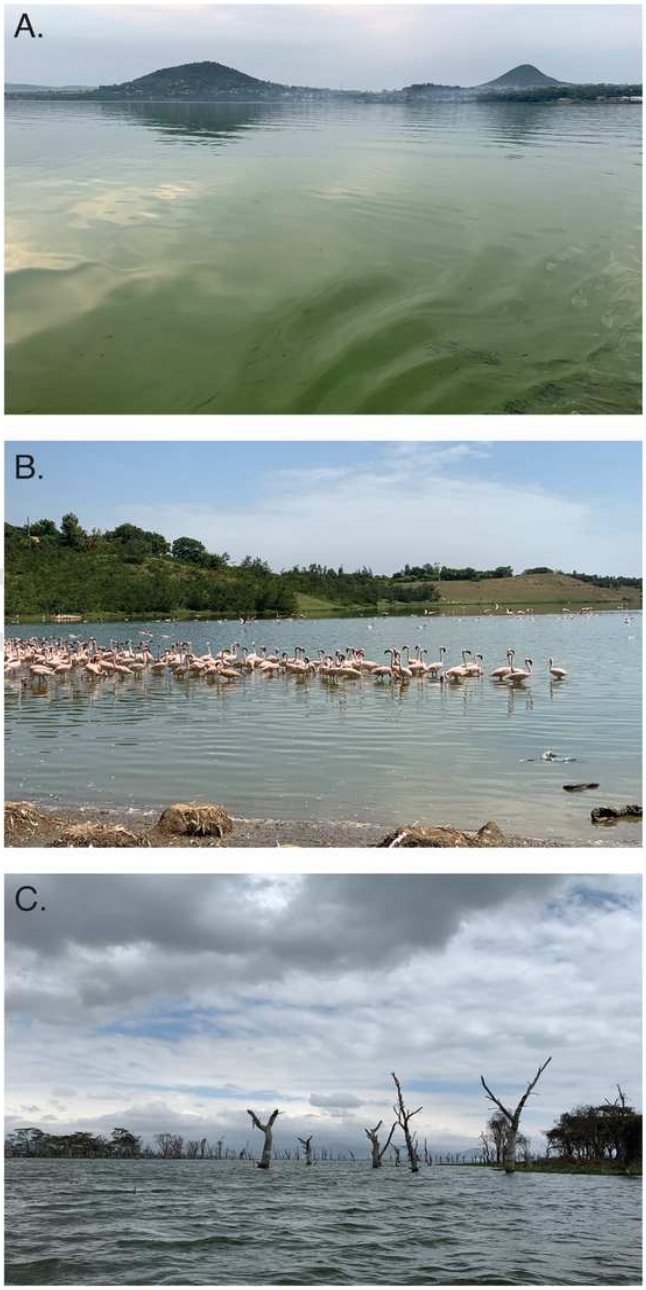

Fig. 1. The three lakes sampled over the course of the 2022 *NSF IRES Advanced Studies Institute on Water Quality and Harmful Algal Blooms in Lake Victoria, Kenya*. (A) Image of a cyanobacterial bloom in Homa Bay, Lake Victoria taken on June 24th, 2022, prior to sampling. (B) Image of flamingoes along the shoreline of Lake Simbi taken on July 2nd, 2022, prior to sampling. (C) Image of Lake Naivasha shoreline taken on July 12th, 2022, prior to sampling. Photo credit: Brittany N. Zepernick.

2.2. DNA extraction and sequencing

Prior to extraction, DNA/RNA Shield was flushed from filters using ~50 mL of sterilized PBS solution to maximize DNA yields (Supplemental Figure 1). DNA was extracted using standard basic phenol-

chloroform methods followed by ethanol precipitation as reported previously (Martin and Wilhelm, 2020; Zepernick et al., 2022). DNA was assessed for quality using a Nanodrop ND-100 Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) to ensure purity was sufficient for sequencing (260/280 > 1.80). DNA concentration was quantified using the Qubit dsDNA HS assay (Invitrogen, Waltham, MA, USA). Libraries for 16S rRNA metabarcoding were prepared using the Nextera XT Library Preparation Kit (Illumina, San Diego, CA, USA) and the dual-index approach targeting the V4 region with the 515F and 806R primers (Gohl et al., 2016; Illumina, 2015). Following, 16S rRNA gene libraries were sequenced on an Illumina MiSeq using the Nano v2 kit (2 × 300 bp, paired end reads) at the *University of Minnesota Genomics Center*. Concurrently, metagenomes were generated (2 × 150 bp paired end reads) on a NovaSeq S4 yielding ~200 million total reads per sample.

2.3. Metabarcoding analysis to identify major cyanobacterial genera

All raw fastq files (n = 26) were processed in RStudio (v.4.3.2) using the DADA2 pipeline (v.1.28.0) (Callahan et al., 2016). Briefly, primers were trimmed from samples with resulting sequence length truncated at 225 bp (forward reads) and 200 bp (reverse reads) based on sequence quality profiles. Quality control was performed using the DADA2 error correction model, with sequences possessing > 2 expected errors removed and sequences with quality scores < 2 truncated. Forward and reverse reads were interleaved with a minimum overlap of 150 bp and chimeras, chloroplasts and mitochondria were removed from the dataset. Subsequently, taxonomy was assigned to amplicon sequencing variants (ASVs) using the DADA2 classifier and the DADA2 formatted SILVA ribosomal RNA gene database (v.138.1) (Quast et al., 2012). Assignments were made to the species level at 100 % identity for ASVs and taxonomically unassigned chloroplasts were removed. The resulting ASV count matrix (n = 4,383 unique ASVs) was used for downstream statistical analysis (Supplemental Table 2) with results from the six libraries of interest reported in this study.

2.4. Metagenomic analysis reveals minor cyanobacterial genera and metabolic potential

All raw fastq files (n = 26) were preprocessed (quality control, filtering, trimming) using fastp (v.0.20.0) (Chen, 2023; Chen et al., 2018) through the Great Lakes Atlas of Multi-omics Research (GLAMR) pipeline (https://greatlakesomics.org/). Files were subsequently processed on the ISAAC NG cluster. Libraries were concatenated and assembled (coassembled) using MEGAHIT (v.1.2.9) (Li et al., 2015) with coassembly statistics determined *via* QUAST QC (Supplemental Table 3) (v.5.0.2) (Gurevich et al., 2013). Trimmed reads of each library were mapped to the coassembly using BBMap (default settings) (Supplemental Table 4,5) (v.38.90) (Bushnell, 2014). Open reading frames and gene predictions within the coassembly were called using MetaGeneMark (v.3.38) (Zhu et al., 2010) using the metagenome style model. Following, read counts to the open reading frames were tabulated using featureCounts available in the subread package (v.2.0.2) (Liao et al., 2014). Taxonomic annotations of the prokaryotic community were performed using the nucleic acid sequences corresponding to open reading frames and Kraken2 (v.2.1.3) (Kraken2 database = standard) (Wood et al., 2019). We note that while *Anabaena* was reclassified as *Dolichospermum* spp. (Driscoll et al., 2018; Li et al., 2016), the most recent version of the KRAKEN2 database does not reflect this revision – thus we have kept the annotations of *Anabaena* and *Dolichospermum* spp. separate. Likewise, we recognize *Cylindrospermopsis* spp. was reclassified as *Raphidiopsis* spp. (Aguilera et al., 2018), yet this revision is also not included within the most recent reference databases. Thus, we kept these genera separated within their original classifications for reproducibility. Functional annotations of protein sequences corresponding to the open reading frames were made using eggNOG (v.2.1.12) (Huerta-Cepas et al., 2019) and the eggNOG database (v.5.0.2). All subsequent analyses in this study used the six metagenomic (MEG) libraries of interest.

2.5. Comparative analyses between metabarcoding and metagenomic taxonomic distributions

To our knowledge, microbiota within two of the three lakes in this study have never been characterized by molecular tools – raising concerns dominant genera may be underrepresented or absent within traditional taxonomic databases which are biased towards ecosystems within the global north (Hughes et al., 2021; Wishart, 1998). To determine reproducibility of metabarcoding results and enable genomic potential predictions, metagenomic analysis was conducted in tandem with metabarcoding analysis on the same DNA samples. In further efforts to confirm reproducibility – two common taxonomic reference databases were used: Metabarcoding analyses utilized the taxonomic reference database SILVA and metagenomic analyses used the taxonomic database KRAKEN2. While metabarcoding is a time and cost effective means to determine the major taxa within communities, metagenomics possesses the sequencing depth and power to identify lesser abundant, minor taxa which remain biologically meaningful (Durazzi et al., 2021). In the present study, we employ metabarcoding to identify major cyanobacterial genera (*i.e.,* bloom formers) and trial its effectiveness at capturing community abundance and diversity as a potential cost-effective monitoring strategy for local agencies. In tandem, we employ metagenomics to confirm major genera identified *via* metabarcoding, investigate minor cyanobacterial genera and identify the biosynthetic potential of both major and minor genera with respect to toxin production.

2.6. Cyanotoxin and cyanopeptide gene cluster detection

Trimmed, quality-controlled metagenomic libraires (n = 6) were mapped to biosynthetic gene clusters (BGCs) in the MIBiG database using MiniMap2 (v.2.24) (Li, 2018). Briefly, reads were mapped to BGCs that were taxonomically annotated as toxigenic cyanobacterial genera including *Anabaena, Aphanizomenon, Raphidiopsis, Microcystis, Dolichospermum, Planktothrix, and Sphaerospermopsis* spp. Results were statistically filtered to include mappings with 80 % coverage and 90 % identity to the reference. Hits to genes within BCGs were retained if ≥ 50 % of the gene was covered in the library and ≥ 75 % of core genes within the BGC were present. Subsequently, read counts were normalized to mean reads per kilobase per million (RPKM) of core genes present for downstream analysis (Wagner et al., 2012).

2.7. Calculating cHAB risk and proposing a monitoring framework

Ten risk metrics pertaining to the formation of cHABs (*i.e.,* trophic status, water temperature, residence time, wet / dry season, watershed land use, riverine inputs, lake depth) and human exposure to cHABs (*i.e.,* population in watershed, potable water source, recreational resource) were selected based on the literature (Olokotum et al., 2020; Omara et al., 2023; Veerman et al., 2024). Following, each risk metric was assigned point designations: Trophic status: Oligotrophic = 0, mesotrophic = 1, eutrophic = 2, hypereutrophic = 3, water temperature: <15 °C = 0, 16-23°C = 1, 24-30 °C = 2, > 30 °C = 3, residence time: <1yr = 0, 1-3yrs = 1, 3-6yrs = 2, >6yrs = 3, wet/dry season: wet = 0, dry = 1, watershed land use: unpopulated = 0, urban = 1, agriculture small family farms = 2, agriculture large industry farms = 3, agriculture livestock = 3, riverine inputs: 1 point assigned per river draining into lake, lake depth: > 100 m = 0, 100-75 m = 1, 75-10 m = 2, <10 m = 3, population in watershed: <300,000 = 1,

300,000-800,000 = 2, 800,000-1.5 million =3, >1.5 million = 4, potable water source: No = 1 Yes = 3, recreational use: no = 1, yes = 2. Subsequently, scores for each risk metric were assigned per lake based on the water column physiochemistry at the time of sampling (Bullerjahn et al., 2024) and current knowledge of lake features (Table 1). The sum of the ten cHAB risk metrics was then calculated, leading to a risk score designation of low risk = < 15, moderate risk = 16-24, or high risk = > 25. In addition to cHAB risk estimates, a potential monitoring framework was formulated based on the results of this study and the recommendations of local stakeholders. Subsequently, we provide a user-friendly excel sheet which local researchers can use to calculate risk scores for these systems.

2.8. Statistical analysis

Normalization of raw read mappings to the metagenomic coassembly (RPKM) (Dick, 2018) was performed on the ISAAC NG OnDemand RStudio Server (v.4.0.4). Calculations of ASV alpha diversity metrics (species richness, evenness, Shannon's H) and MEG beta diversity metrics (Bray-Curtis dissimilarity and nonmetric multidimensional scaling - NMDS) were also performed on the RStudio Server. Similarity Percentages (SIMPER) were calculated in PRIMER (v.7) (Clarke and Gorley, 2015). KEGG Orthology (Kanehisa et al., 2016) gene set enrichment analyses were performed using clusterProfiler (v4.10.1) (Settings: pAdjustMethod = Benjamini & Hochberg, minGSSize = 10, seed = 123) (Yu et al., 2012) on the RStudio Server. Heatmaps were made using heatmapper.ca (clustering method = Average linkage, distance measurement method = Euclidean) (Babicki et al., 2016). Comparisons of dissolved nutrients and alpha diversity across lakes were performed using one-way ANOVAs with multiple comparisons *via* Prism (v.10.0.3).

3. Results

3.1. Distinct biotic and physiochemical profiles distinguish lakes

Cyanobacterial blooms were visually observed in Lakes Victoria and Simbi at the time of sampling. In Lake Victoria's Winam Gulf, expansive green surface scums were spotted (Fig. 1A) with Algae Torch readings estimating cyanobacteria accounted for ~60 % of the total photosynthetic community (Total Chl *a* = 75 μg • L^{-1}) (Supplemental Table 1). Likewise, Lake Simbi was dominated by green filamentous aggregates which were confirmed to be *Limnospira* spp. by brightfield microscopy (Fig. 1B) (Supplemental Figure 2). In contrast, there was no qualitative nor quantitative evidence of a bloom at Lake Naivasha at the time of sampling (Fig. 1C). Regarding nutrient profiles, Lake Simbi exhibited nitrogen (NO_x, NH_4) and phosphorus (DRP) concentrations that were significantly higher compared to the Winam Gulf and Lake Naivasha ($p \leq 0.008$) (Supplemental Figure 3) (Supplemental Table 1). Yet, higher (albeit non-significant, $p \geq 0.05$) concentrations of DRP (7.62 μg • L^{-1} +/- 4.64 compared to 0.22 μg • L^{-1} +/-0.31), NO_3 + NO_2 (1.26 μg • L^{-1} +/- 0.76 compared to 0.25 μg • L^{-1} +/- 0.26) and NH_4 (27.91 μg • L^{-1} +/- 19.52 compared to 7.10 μg • L^{-1} +/- 5.09) were recorded in the Winam Gulf compared to Lake Naivasha.

3.2. Metabarcoding confirms blooms of major cyanobacterial genera in the Winam Gulf and Lake Simbi

While each lake had comparable percentages of cyanobacteria at the phylum level (23-35 %) (Fig. 2A) (Supplemental Table 6), distinct differences in the dominant cyanobacterial genera were present (Fig. 2B) (Supplemental Table 7). The Winam Gulf site was dominated by *Dolichospermum* spp. (83 % +/- 0.84) and a subpopulation of *Microcystis* spp. (14 % +/- 0.54). In contrast, Lake Simbi was dominated by *Limnospira* spp. (76 % +/- 10.27) with a subpopulation of *Nostoc* spp. (24 % +/- 10.11). The cyanobacterial community of Lake Naivasha was largely composed of *Cyanobium* spp. (58 % +/- 0.48) with a subpopulation of *Microcystis* spp. (2 % +/- 0.07). We note *Cyanobium* is a routinely abundant, non-toxic member of many comparatively healthy freshwater lakes (Ivanikova et al., 2008; Wilhelm et al., 2006) and thus this is not considered a bloom – which we define as one genus contributing > 55 % of the photosynthetic community which alters the normal ecological function of the system (Zepernick et al., 2024).

Notably, while the cyanobacterial community was largely dominated by one or two genera in Winam Gulf and Lake Simbi (consistent with a bloom), Lake Naivasha harbored a more diverse cyanobacterial community (Fig. 2C) with significantly higher alpha diversity metrics ($p \leq 0.004$) (Supplemental Figure 4). Specifically, ~40 % of the cyanobacteria in Lake Naivasha constituted <5 % of the total cyanobacterial ASV counts across all three lakes - and were thus classified as "Other" (Fig. 2C). *Synechocystis* spp. and *Raphidiopsis* spp. were the second and third most abundant cyanobacterial genera in Lake Naivasha (each constituting 15 % of the cyanobacterial population), while they were not detected in Homa Bay (Winam Gulf) or Lake Simbi *via* metabarcoding. Likewise, populations of toxin producing *Limnolyngbya* spp. (1 %) and *Planktothrix* spp. (0.5 %) were present in Lake Naivasha but not the others (Supplemental Table 7).

3.3. Metagenomics reveals Simbi and Naivasha harbor diverse, minor populations of cHAB genera

Relative abundances of the major cyanobacterial genera within each of the three lakes were comparable in metabarcoding and metagenomic analyses (Supplemental Figure 5) (Supplemental Table 8) with the top four major cyanobacterial genera detected by each respective analysis weakly correlated ($y = 395.0x + 414{,}539$, $R^2 = 0.40$) (Supplemental Figure 6). Yet, while both analyses confirmed *Limnospira, Dolichospermum, Microcystis* and small picoplankton (*Cyanobium* or *Synechococcus*) were the four major cyanobacterial genera present across the three lakes – metagenomic analyses indicated a higher amount of variability and diversity in the "minor contributors" (*i.e.,* those contributing less than 5 % of the total relative abundance across all sample sites) (Supplemental Figure 7). Metagenomic analyses of the minor contributing genera revealed metabarcoding failed to detect subdominant populations of potentially toxigenic cyanobacteria in Lake Simbi. Notably, ~5 % of the cyanobacterial community in Lake Simbi was annotated as *Raphidiopsis, Dolichospermum, Leptolyngbia, Nodularia, Planktothrix,* and *Sphaerospermopsis* (Supplemental Table 9). Whereas metabarcoding successfully identified these populations within Lake Naivasha –certain genera were underrepresented in metabarcoding results (specifically *Microcystis* abundance).

3.4. Key genera and genes drive dissimilarity between cyanobacterial communities

Preliminary determination of cyanobacterial community similarity between lakes was performed using normalized metagenomic libraries and Similarity Percentages (SIMPER). All three lakes demonstrated high levels of similarity between biological replicates (71–88 % similarity) (Fig. 3A) (Supplemental Table 10). Lake Victoria's Winam Gulf and Lake Simbi demonstrated an average similarity of 47 % whereas the Winam Gulf exhibited 38 % similarity to Lake Naivasha. In turn, Lakes Simbi and Naivasha demonstrated 33 % similarity. All three lakes shared 35 % similarity to one another. Subsequent visualization of beta diversity metrics (Bray-Curtis dissimilarity) *via* NMDS confirmed replicates clearly clustered by lake, with Lake Naivasha libraries clustering away from the Winam Gulf and Lake Simbi on both planes (Fig. 3B) (Supplemental Figure 8).

To determine key genera and metabolic processes driving dissimilarity between cyanobacterial communities, SIMPER genes were assessed for taxonomy and function. Comparisons between the Winam

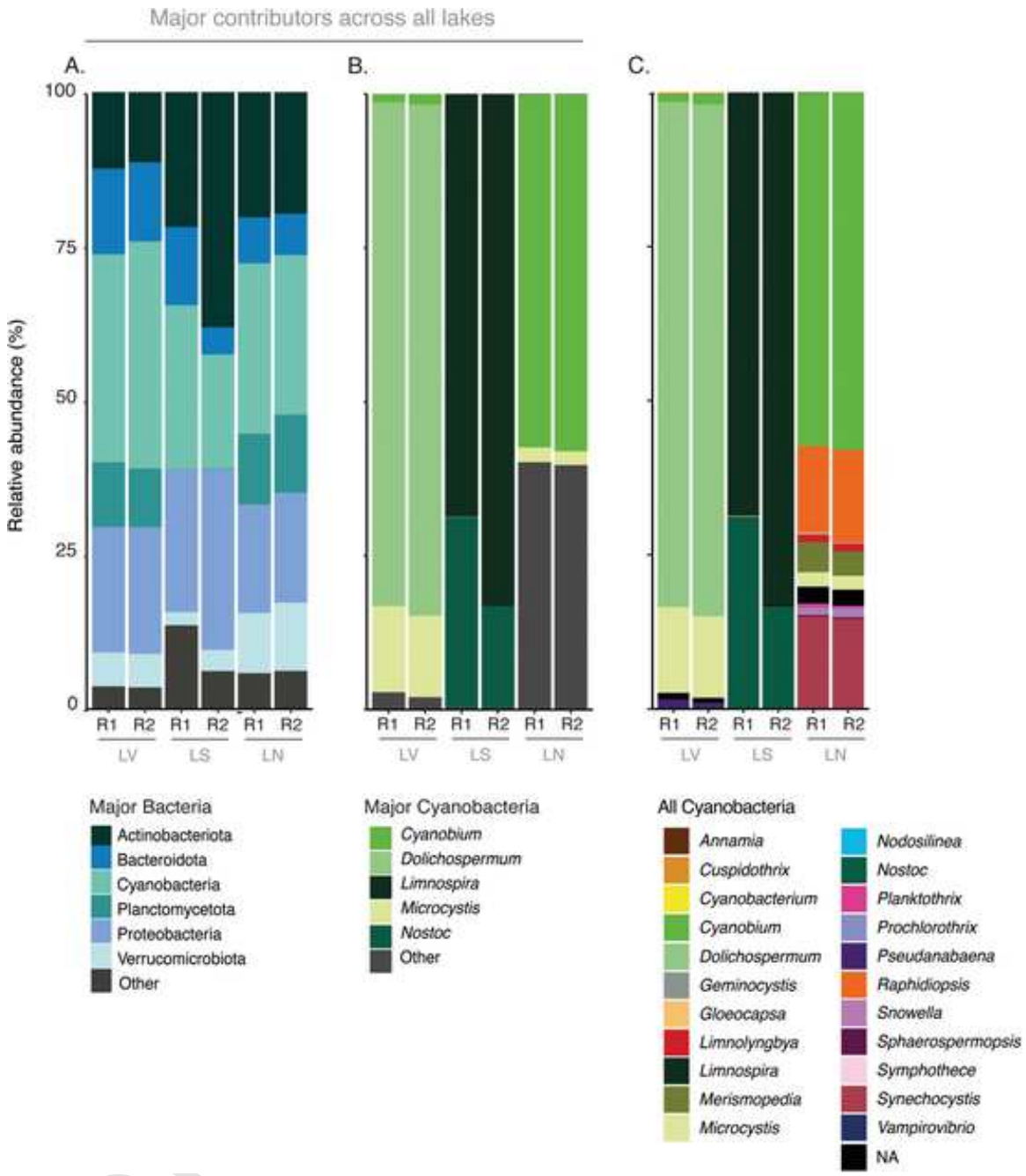

Fig. 2. Metabarcoding relative abundances of major bacterial and cyanobacterial communities (ASVs) across Lake Victoria (LV), Simbi (LS) and Naivasha (LN). (A) Major bacterial phyla abundance across the three lakes. All detected groups which formed < 5 % of the total reads across the three lakes are included within "Other" (*10bav-F6, Acidobacteriota, AncK6, Armatimonadota, Bdellovibrionota, Caldisericota, Campylobacterota, Chloroflexi, Chrysiogenetota, Cloacimonadota, Deferrisomatota, Deinococcota, Dependentiae, Desulfobacterota, Elusimicrobiota, Fibrobacterota, Firmicutes, Fusobacteriota, Gemmatimonadota, Hydrogenedentes, Latescibacterota, LCP-89, Margulisbacteria, MBNT15, Methylomirabilota, Myxococcota, NB1-j, Nitrospinota, Nitrospirota, Patescibacteria, PAUC34f, RCP2-54, SAR324 clade, Spirochaetota, Sumerlaeota, Sva0485, Synergistota, TA06, Thermotogota, WOR-1, WPS-2, WS1*). (B) Major cyanobacterial genera abundance across the three lakes. All detected groups which formed < 5 % of the total reads across the three lakes are included within "Other" (*Annamia, Cuspidothrix, Cyanobacterium, Raphidiopsis, Geminocystis, Gloeocapsa, Limnolyngbya, Merismopedia, Nodosilinea, Planktothrix, Prochlorothrix, Psuedoanabaena, Snowella, Spaerospermopsis, Synechocystis, Vampirovirbio*). (C) Unfiltered cyanobacterial abundance across the 3 lakes. Biological duplicates are joined with horizontal bars.

Gulf and Lake Simbi revealed 1,125 genes formed the top 10 % dissimilarity, with the Gulf *vs.* Lake Naivasha comparisons containing 2,576 genes and Lake Simbi *vs.* Lake Naivasha containing 3,294 genes. With respect to taxonomy, ~53 % and ~57 % of the top 10 % genes driving dissimilarity between Winam Gulf *vs.* Lake Simbi and Winam Gulf *vs.* Lake Naivasha were annotated as *Dolichospermum* and *Limnospira* spp. (Supplemental Figure 9A, B) (Supplemental Tables 11, 12). In comparison, ~92 % of the genes driving dissimilarity between Lakes Simbi *vs.* Naivasha were annotated as *Limnospira* and *Synechococcus* spp. (Supplemental Figure 9C) (Supplemental Table 13). With respect to genomic potential, the clusters of orthologous groups (COGs) contributing to the top 10 % dissimilarity were largely the same across lakes – with genes

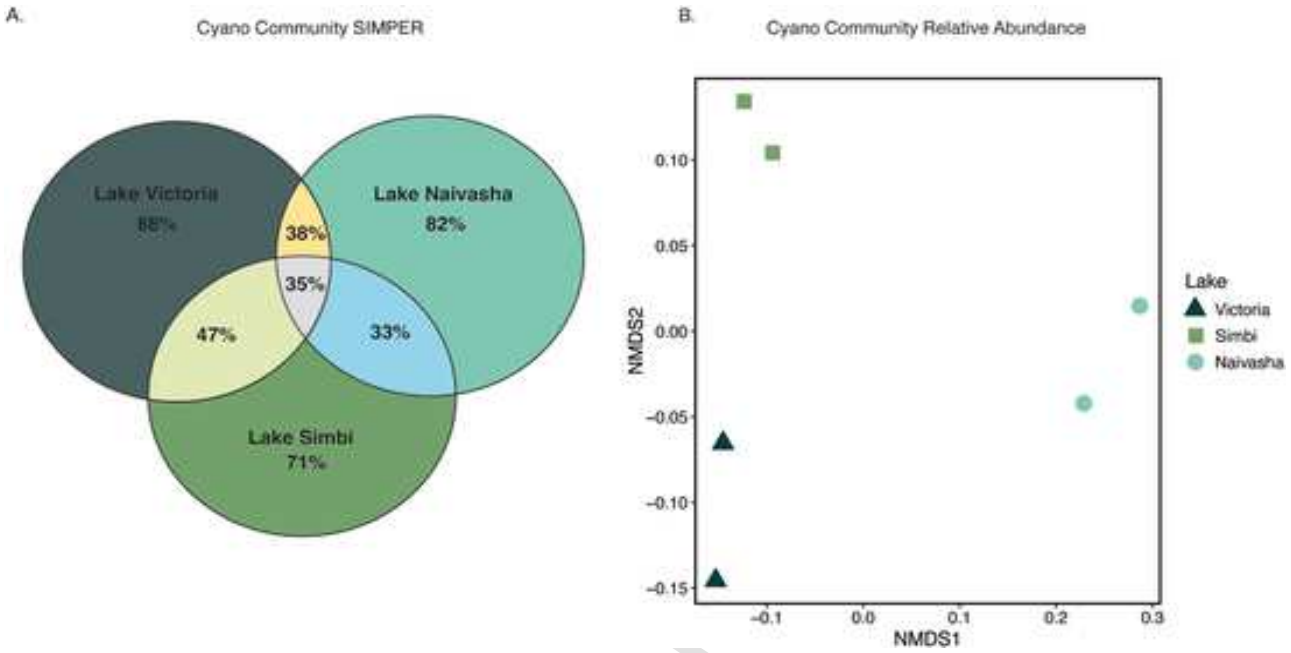

Fig. 3. Beta diversity statistics (Bray-Curtis dissimilarity) of the cyanobacterial communities within the six normalized metagenomic libraries (RPKM). (A) Venn diagram depicting Similarity Percentages (SIMPER) of intra and inter-similarity comparisons of cyanobacterial community. Lake Victoria (dark blue) demonstrated 88 % similarity between biological replicates, 47 % similarity with Lake Simbi (light green), and 38 % similarity with Lake Naivasha (yellow). Lake Simbi (dark green) demonstrated 71 % similarity between biological replicates and 33 % similarity with Lake Naivasha (light blue). Lake Naivasha (teal) demonstrated 82 % shared similarity between replicates. All three lakes demonstrated 35 % shared similarity. (B) NMDS of the cyanobacterial community relative abundance, stress value < 0.05. Lake Victoria is indicated by dark blue triangles (LV), Lake Simbi is indicated by green squares (LS) and Lake Naivasha is indicated by teal circles (LN).

involved in cellular metabolism driving dissimilarity (Supplemental Figure 9D, E, F) (Supplemental Tables 14, 15, 16).

3.5. Gene enrichment analyses indicate differential metabolic processes drive genomic diversity

Gene enrichment analyses were made using normalized metagenomic libraries to determine which core metabolic pathways were driving cyanobacterial community trends at the time of sampling. Lake Simbi analyses lacked statistical significance and yielded no results, which was attributed to the dominance of a single genera within the water column (*Limnospira* spp.). Thus, only the Winam Gulf and Lake Naivasha libraries were pursued for downstream analysis. In total, 18 KEGG pathways were statistically enriched within the Winam Gulf and 89 pathways were significantly enriched in Lake Naivasha ($p_{adj} < 0.05$) (Supplemental Tables 17, 18). Comparisons between the top 20 enriched pathways in Lake Naivasha *vs.* the Winam Gulf indicated four distinct pathways of interest (Fig. 4). Cysteine and methionine metabolism (ko00270) was the 10th highest enriched pathway in the Winam Gulf yet failed to make the top 20 in Lake Naivasha (Fig. 4A). In contrast, homologous recombination (ko03440), 2-oxocarboxylic acid metabolism (ko01210) and methane metabolism (ko00680) were in the top 20 pathways enriched in Lake Naivasha, yet these pathways were not statistically enriched in the Winam Gulf.

3.6. Cysteine metabolism and sulfur-scavenging genes are associated with a diazotrophic bloom in Lake Victoria's Winam Gulf

Considering the gene enrichment results, we investigated how metabolic potential may drive cyanobacterial community diversity between a cHAB-dominated site in the Winam Gulf and cHAB subdominated Lake Naivasha. The enrichment of sulfur-heavy amino acid metabolism in the Winam Gulf (Fig. 4A) prompted an investigation into sulfur-related genes within Winam Gulf *vs.* Lake Naivasha SIMPER results (Supplemental Table 15) – which revealed 53 genes related to sulfur transport and metabolism were enriched in abundance within Winam Gulf (Supplemental Figure 10A) (Supplemental Table 19). Sulfate/sulfonate transporters (*cysA, cysW,* and *cysT*), genes encoding for sulfur-heavy photosynthetic proteins (*ndhD, ndhJ, ndhK, ndhI, ndhN, ndhH, ndhA, ndhA, ndhB, ndhC, ndhM, ndhO*) and genes involved in iron-sulfur binding and repair (*patB, nifU_1, nifU_2, fdxB, hycB, mrp*) were present. Sulfur metabolism genes belonged to filamentous, diazotrophic cyanobacteria, including the major bloom-former in the Winam Gulf (*Dolichospermum* spp.) in addition to *Nostoc, Anabaena* and *Sphaerospermopsis* spp. (Supplemental Figure 10B). Coincident with gene enrichment results, ~34 % of genes involved in sulfur scavenging and recycling contributed to SIMPER dissimilarity between Winam Gulf and Lake Naivasha – including genes annotated as cysteine desulfurases (*cdsF_1, cdsF_2, iscS_1, iscS_2*), genes encoding for sulfatases (*smfl_1, smfl_2, smfl_3, smfl_4, smfl_5, smfl_6*) and nucleotide-disulfide oxidoreductases (*selD, gor, merA*). Notably, *Dolichospermum* spp. possessed three of the four genes in the *nif* system which has been shown to direct nitrogenase maturation by coordinating the biogenesis of Fe/S proteins (Supplemental Figure 10C) (Blanc et al., 2015; Frazzon and Dean, 2003; Jacobson et al., 1989).

3.7. Methane metabolism genes were significantly enriched in Lake Naivasha's picocyanobacteria

Considering the significantly enriched pathways of methane metabolism and 2-oxocarboxylic acid in Lake Naivasha (Fig. 4B), we assessed a variety of methane-related metabolic processes alluded to within the Winam Gulf *vs.* Lake Naivasha SIMPER results (Supplemental Table 15). Notably, genes encoding for methylphosphonate transporters (*phn*)

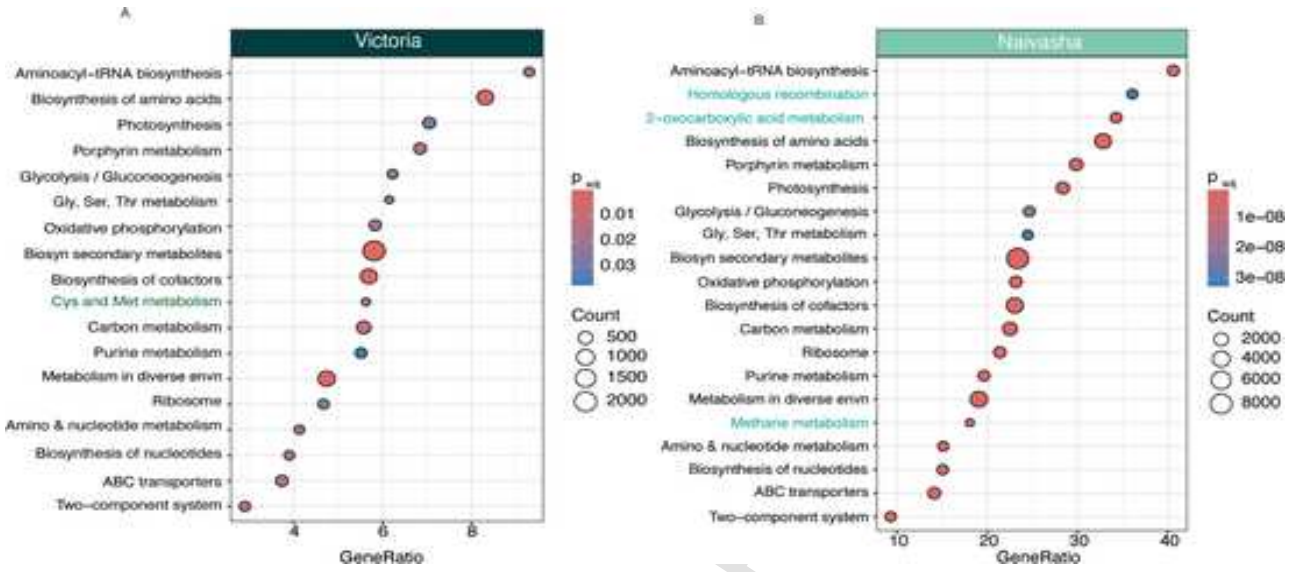

Fig. 4. Gene set enrichment analyses in Lake Victoria (dark blue) and Naivasha (teal) organized by KEGG mapper metabolic pathways. (A) Gene enrichment results showing the 18 significantly enriched metabolic pathways in Lake Victoria organized by gene ratio and significance. Metabolic pathways uniquely enriched in Lake Victoria (*i.e.,* not found in the top 20 significantly enriched pathways in Naivasha) are shown in dark green. (B) Gene enrichment results showing the top 20 significantly enriched metabolic pathways in Lake Naivasha organized by gene ratio and significance. Metabolic pathways uniquely enriched in Lake Naivasha (*i.e.,* not found in the 18 significantly enriched pathways in Victoria) are shown in light blue. The number of normalized read counts per KEGG mapper pathway is indicated by circle size.

were found to drive dissimilarity between the two lakes. Further analysis identified 213 phosphonate genes within the coassembly which were nearly unanimously increased in relative abundance within Lake Naivasha (Fig. 5A) (Supplemental Table 20). Taxonomic analysis revealed ~77 % of *phn* genes belonged to non-toxic picoplankton which dominated Lake Naivasha's cyanobacterial community (*Cyanobium* and *Synechococcus* spp.) (Fig. 5B). Yet, despite only ~4 % of all *phn* genes belonging to *Raphidiopsis* spp., it was the only genus that possessed *phnI, phnJ, phnK*, and *phnL* (Fig. 5C), which are phosphonate genes associated with the metabolic conversion of methylphosphonate into methane and phosphate during phosphate limitation (Fig. 5D) (Kamat et al., 2011). In contrast, *Cyanobium* and *Synechococcus* spp. only possessed phosphonate transporters (*phnC, phnD* and *phnE*). Beyond phosphonate transport / metabolism genes, we queried our coassembly against the KEGG methane metabolism pathway (map00680), which revealed a variety of genes associated with methane metabolism (Fig. 6) (Supplemental Table 21, 22). Notably, ~73 % of genes encoding for proteins in the methanotroph serine pathway (KEGG MD: 00346) belonged to *Synechococcus* spp. (Fig. 6A), with these genes increased in relative abundance within Lake Naivasha and decreased in Winam Gulf and Lake Simbi (Fig. 6B). Likewise, ~64 % of genes encoding for proteins involved in the methanotroph ribulose monophosphate pathway (KEGG MD: 00345) belong to *Synechococcus* and *Cyanobium* spp. (Fig. 6C), with these genes increased in relative abundance within Lake Naivasha compared to the others (Fig. 6D). We note while formaldehyde assimilation is the second step of the methane metabolism pathway (as methane is oxidized into formaldehyde), methane transporters and primary genes involved in the first step of methane oxidation (*i.e.,* methane monooxygenase) were not detected.

3.8. Homologous recombination was significantly enriched in Lake Naivasha Synechococcus spp.

The KEGG homologous recombination pathway (KEGG ko03440) was significantly enriched in Lake Naivasha (Fig. 4B) but not in the Winam Gulf (Fig. 4A). We detected 1,441 genes encoding for proteins involved in homologous recombination which largely belonged to *Synechococcus* spp. (~75 %) and *Cyanobium* spp. (~11 %) (Supplemental Figure 11A). Overall, genes involved in homologous recombination were increased in relative abundance within Lake Naivasha and decreased in the others (Supplemental Figure 11B) (Supplemental Table 23).

3.9. Elevated genomic potential for cyanotoxins in Lake Victoria's Winam Gulf and Lake Naivasha

Several BGCs encoding for cyanotoxins were detected in metagenomic samples collected from the Winam Gulf and Lake Naivasha (Fig. 7) (Supplemental Table 24). In contrast, cyanobacterial BGCs were not detected in Lake Simbi replicates. BGCs encoding for secondary metabolites synthesized by *Microcystis* spp. dominated both lakes, with the microcystin BGC (*i.e., mcy* BGC) detected in both Winam Gulf replicates (Fig. 7A), yet neither Lake Naivasha replicate (Fig. 7B). Additional *Microcystis*-derived BGCs (piricyclamide, micropeptin, cyanopeptolin and aeruginosin) were also detected in both lakes. Notably, the anacyclamide BGC belonging to *Anabaena* spp. dominated the Winam Gulf BGC pool (Fig. 7A) yet was not detected in Lake Naivasha (Fig. 7B). In contrast, *Planktothrix*-synthesized anabaenopeptin and *Microcystis*-synthesized microginin were abundant in Lake Naivasha samples yet absent from the Winam Gulf samples. The *Sphaerospermopsis*-synthesized anacyclamide BGC was detected in the Gulf but not Lake Naivasha.

3.10. cHAB risks prevalent in Lake Naivasha and Lake Victoria's Winam Gulf

cHAB risk scores (accounting for likelihood of cHAB presence and human exposure) were calculated for each lake using ten selected risk metrics based on prior regional studies (Supplemental Table 25) (Olokotum et al., 2020; Sitoki et al., 2012; Veerman et al., 2024). Cumulatively, the Winam Gulf was assigned a risk score of 27, placing its cHAB risk as "high" (high = risk score > 25) (Fig. 8A). Lake Naivasha

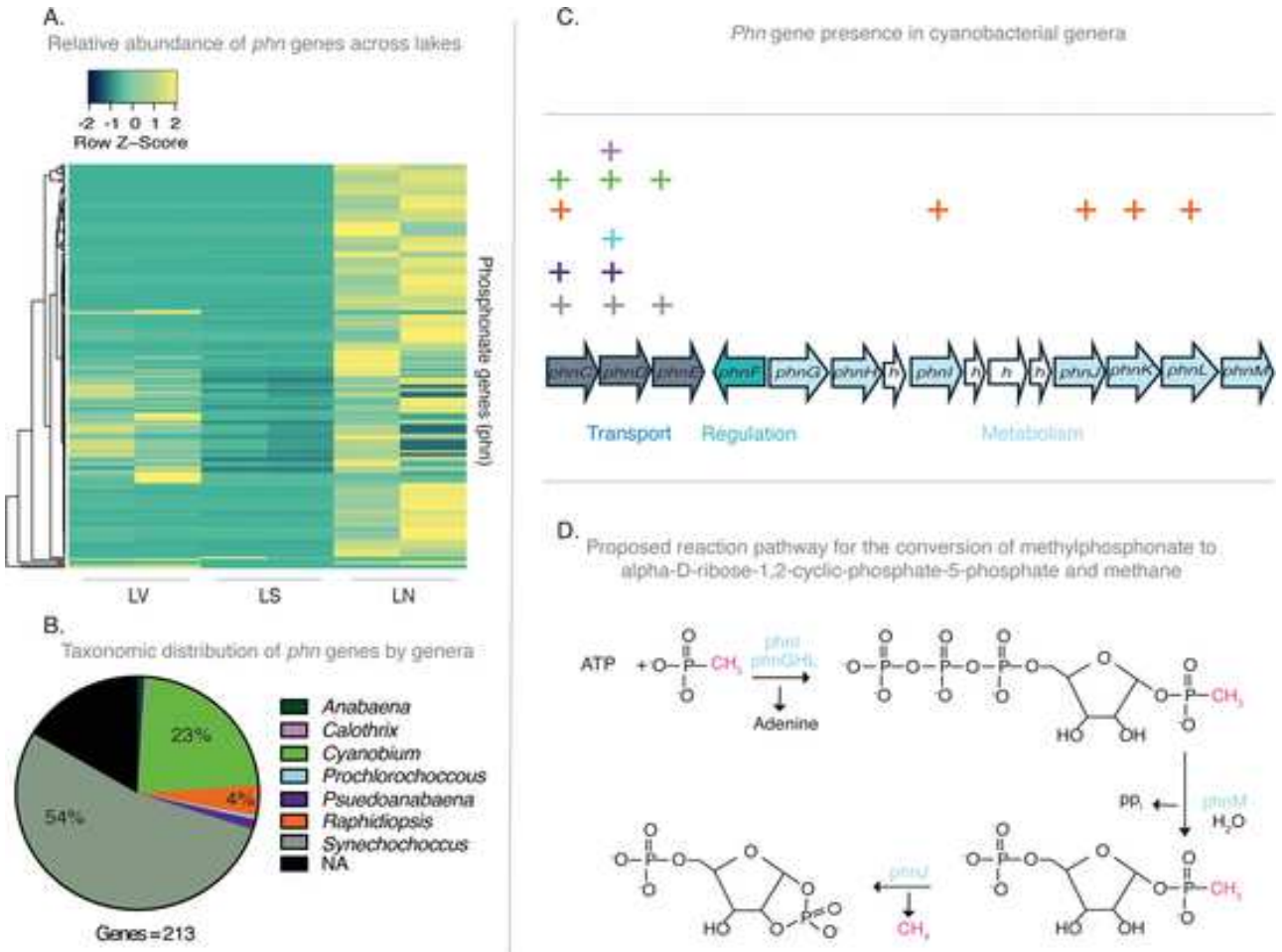

Fig. 5. Taxonomic and functional analysis of phosphonate genes (*phn*) across Lake Victoria (LV), Simbi (LS) and Naivasha (LN). (A) Heat map depicting the relative abundance (RPKM) of the 213 *phn* genes within the coassembly. (B) Taxonomic distribution of the 213 *phn* genes in the coassembly by cyanobacterial genera: *Anabaena* spp. (green), *Calothrix* spp. (lavender), *Cyanobium* spp. (bright green), *Prochlorococcus-like* spp. (light blue), *Psuedanabaena* spp. (dark purple), *Raphidiopsis* spp. (orange), *Synechococcus* spp. (sea foam green), not annotated (NA, black). (C) Presence / absence of the different *phn* genes within the phosphonate operon in major cyanobacterial genera which are color coded in the same manner as panel B. Genes *phnC, phnD,* and *phnE* are involved in intracellular phosphonate transport (dark blue), *phnF* regulates expression of genes in the phosphonate operon (turquoise), genes *phnG, phnH, phnI, phnJ,phnK, phnL* and *phnM* are thought to be involved in phosphonate metabolism (light blue). Hypothetical genes and their placements are predicted by a prior study (white "h" arrows). (D) Proposed reaction pathway for the conversion of methylphosphonate to alpha-D-ribose-1,2,-cyclic-phosphate-5-phosphate and methane.

scored moderate (score: 23, moderate = risk score 16-24) and Lake Simbi scored low (score: 14, low = risk score < 15) (Fig. 8A). Discussions amongst student *NSF-IRES* participants were aligned with the results of this study – culminating into a targeted PCR screening proposal for local academic and government agencies to implement (Fig. 8B). Briefly, the cyanotoxin BGCs used in this study have been provided (Supplemental Table 26) to identify potential primer candidates of interest for future implementation by local researchers, lake managers and municipal authorities. By providing the full nucleotide sequences of the genes present in cyanotoxin BGCs which were deemed present (or could be present) in the sample sites, primers can be optimized for future, regional use in these understudied, tropical lakes. In turn, this approach may circumvent limitations of traditional, commercially available kits such as the Phytoxigene test (Phytoxigene Inc., Akron, OH, USA) - which only selects for a single gene in the cyanotoxin BGC (*e.g.* Microcystin - *mcyE*, cylindrospermopsin - *cyrA*) and may be biased towards cyanotoxin genes present in well-studied, temperate systems.

4. Discussion

Historically, cHAB studies have been biased towards the global north (Lawrence et al., 2023; Svirčev et al., 2019). Indeed, the majority of *Microcystis* spp. public health guidelines are based on a few temperate systems (Roegner et al., 2020). Yet, *Microcystis* spp. is globally distributed with blooms documented in > 108 countries (Harke et al., 2016). This lack of sample size and diversity (especially with respect to tropical freshwater systems and the global south) confers a multitude of disadvantages to the field and threatens the health of millions of Africans who depend on these vital resources (Olokotum et al., 2020). In the present study, we employed DNA sequencing to establish baseline, snapshot molecular characterizations of cyanobacterial communities in three anthropogenically-important and ecologically distinct Kenyan lakes.

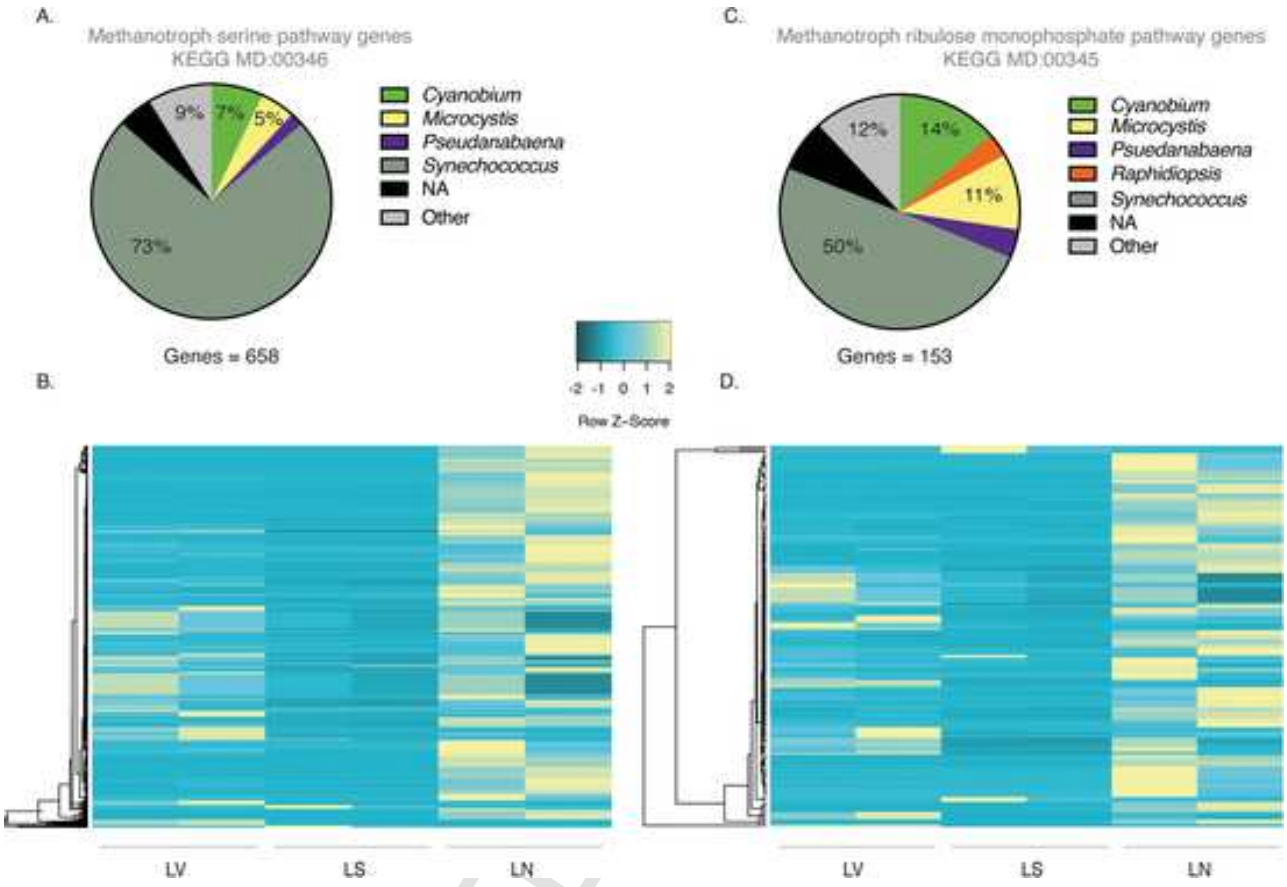

Fig. 6. Taxonomic and functional analysis of methanotrophic genes across Lake Victoria (LV), Simbi (LS) and Naivasha (LN). (A) Taxonomic distribution of the 658 genes annotated in the methanotrophic serine pathway (KEGG MG:00346) in the coassembly by cyanobacterial genera: *Cyanobium* spp. (bright green), *Microcystis* spp. (yellow), *Psuedanabaena* spp. (dark purple), *Synechococcus* spp. (sea foam green), not annotated (NA, black), Other (grey). "Other" contains genes that contribute <1 % abundance (RPKM) across all lake replicates. Other: *Acarychloris, Anabaena, Calothrix, Chroococcidiopsis, Crocosphaera, Cyanoabcterium, Cylindrospermopsis (Raphidiopis), Dolichospermum, Fischerella, Geitlerinema, Gloeothece, Halothece, Halotia, Kovacikia, Leptodesmis, Limnospira, Moorena, Nodosilinea, Nodularia, Nostoc, Oxynema, Parasynechococcus, Phormidium, Picosynechococcus, Richelia, Rippkaea, Sphaerospermopsis, Thermocoleostomius, Thermosynechococcus* and *Trichothermofontia* spp.).(B) Heat map depicting the relative abundance (RPKM) of the 658 serine methanotrophy genes within the coassembly. (C) Taxonomic distribution of the 153 genes annotated in the methanotrophic ribulose monophosphate pathway by cyanobacterial genera: *Cyanobium* spp. (bright green), *Microcystis* spp. (yellow), *Psuedanabaena* spp. (dark purple), *Raphidiopsis* spp. (orange), *Synechococcus* spp. (sea foam green), not annotated (NA, black), Other (grey). "Other" contains genes that contribute <2 % abundance (RPKM) across all lake replicates. Other: *Anabaena, Geitlerinema, Gloeobacter, Gloeocapsopsis, Leptodesmis, Leptolyngbya, Limnospira, Nodularia, Nostoc, Rivularia, Synechoystis* and *Thermostichus* spp.

4.1. Winam Gulf dominated by non-toxic Dolichospermum spp. blooms with evidence of sulfur constraint on growth

The Winam Gulf routinely exhibits shifts in cyanobacterial blooms between filamentous, diazotrophic *Dolichospermum* spp. (Gikuma-Njuru et al., 2013) and colonial, non-diazotrophic *Microcystis* spp. (Sitoki et al., 2012). The Gulf also harbors spatially distinct cyanobacterial communities. Notably, *Raphidiopsis* spp. dominated the eastern regions of the Winam Gulf in June-July of 2022 (Brown et al., 2024a), whereas *Dolichospermum* spp. dominated the southern region of the Gulf (Homa Bay) during the same sample period in the present study. While the drivers that yield these distinct populations remain unclear (Olokotum et al., 2020), prior studies have proposed nitrogen limitation as the major selector for *Dolichospermum* spp. (Gikuma-Njuru and Hecky, 2005; Mugidde et al., 2003). Here, we provide evidence that sulfur metabolism shaped cyanobacterial competition during a *Dolichospermum* spp. bloom in Homa Bay. The synthesis of nitrogenase is energetically expensive, involving the production of enzymes rich in Fe-S clusters (Frazzon and Dean, 2003; Jacobson et al., 1989) which renders them uniquely susceptible to iron and sulfur limitation (Hoffman et al., 2014). Prior studies have shown intracellular sulfide (specifically cysteine) is utilized by cyanobacteria during sulfur limitation (Kharwar et al., 2021; Kumaresan et al., 2017). Notably, cysteine metabolism genes were significantly enriched within Winam Gulf *Dolichospermum* spp. In turn, this increased abundance of sulfate transporters, cysteine desulfurases and sulfur recycling genes suggests diazotrophy in the Winam Gulf *Dolichospermum* spp. bloom could be constrained by sulfur availability.

Beyond the sulfur requirements of nitrogenase, phycobilosomes (and photosynthetic machinery broadly) pose another opportunity for sulfur limitation. A prior study deduced phycobiliproteins play a unique and differential role in *Microcystis* and *Dolichospermum spp.* responses to nitrogen limitation – with *M. aeruginosa* degrading its phycobiliproteins (downregulating photosynthesis) in response to nitrogen limitation and *D. flos-aquae* increasing phycobiliprotein production as nitrogenase-mediated photosynthesis increases (Wang et al., 2021). Cysteine residues play a significant role in phycobiliproteins as they covalently anchor phycobilins and maintain their structural integrity (Kannaujiya et al., 2017; Li et al., 2023). Cumulatively, we hypothesize genes involved in intracellular sulfur scavenging were selected for in *Dolichospermum* spp. within the Winam Gulf.

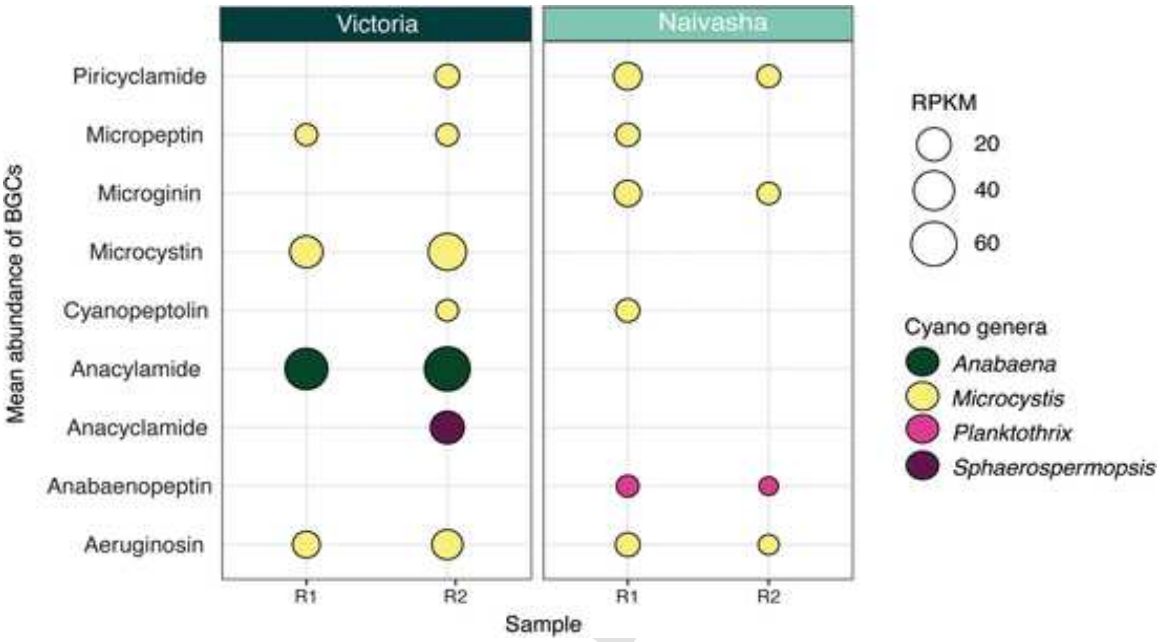

Fig. 7. Cyanopeptide and cyanotoxin biosynthetic gene cluster mean relative abundance (RPKM) across Lake Victoria (dark blue) and Naivasha (teal) replicates (R1, R2). BGCs encoding the biosynthesis of cyanopeptides and cyanotoxins from the MIBiG shown on the y-axis, with the mean normalized abundance (RPKM) of each whole BGC shown as the circle size and the taxonomic designation of the BGC indicated by color: *Anabaena* spp. (dark green), *Microcystis* spp. (yellow), *Planktothrix* spp. (bright pink), and *Sphaerospermopsis* spp. (plum). Note no complete BGCs were detected in Lake Simbi thus it is not included in this figure.

Though sulfur limitation is rarely discussed compared to other nutrients, modern day freshwater lakes (such as Lake Superior) are recognized as low sulfate (< 500 μM) environments (Fakhraee et al., 2017; Phillips et al., 2023). Indeed, lake sulfur concentrations have been reported to vary from 10-1,000 μM (Holmer and Storkholm, 2001; Tipping et al., 1998), with concentrations reported to oscillate on daily and seasonal scales (Norici et al., 2005). Notably, historical studies have routinely described "remarkably low" sulfate concentrations in Lake Victoria of 3-4 μM (Hecky, 1993; Lehman and Branstrator, 1993), with a subsequent study speculating Victoria has "the lowest sulfur concentrations of any large body of water on the planet" (Lehman and Branstrator, 1994). While *in situ* sulfur-uptake experiments deemed these low sulfate levels to be non-limiting to phytoplankton (Lehman and Branstrator, 1994), researchers failed to report the phytoplankton genera in these studies and only performed assays for a 48 h duration. In total, we suggest sulfur may play an underappreciated role in diazotroph *vs.* non-diazotroph cHAB competition within the Winam Gulf.

Subsequently, we investigated the biosynthetic potential of Winam Gulf cHAB genera to produce cyanotoxins. The BGCs detected, including the toxigenic *mcy* BGC, largely belonged to *Microcystis* spp. despite its subdominance. This was consistent with a previous metabarcoding survey of the Winam Gulf which described a *Dolichospermum*-dominated water column during June 2022, yet detected negligible microcystins (Brown et al., 2024a). Building upon this prior observation, we did not detect the *mcy* BGC in *Dolichospermum* spp. across 2022 Winam Gulf MEG samples and Hart et al. (2024 in review) did not detect the *mcy* BGC in *Dolichospermum* spp. during subsequent 2023 MEG sampling efforts. While the *mcy* BGC has been detected in many *Dolichospermum* spp., it remains absent in others (Capelli et al., 2017; Rouhiainen et al., 2004; Yancey et al., 2023). In turn, this invokes questions regarding the biochemical role and competitive trade-off of microcystins given their function in combatting both oxidative (Zilliges et al., 2011) and temperature stress (Stark et al., 2023). In total, our results suggest microcystins in Lake Victoria's Winam Gulf are likely synthesized by *Microcystis* spp.

4.2. cHABs in Lake Naivasha may be outcompeted by methanotrophic picoplankton - evidence for a cyanobacterial methane source and sink

In contrast to the Winam Gulf, there are few reports of cHABs in Lake Naivasha to date. Indeed, we did not find evidence of a cyanobacterial bloom at the time of sampling and molecular investigations indicated ~60 % of cyanobacterial ASVs belonged to routinely abundant, non-toxic picoplankton. Yet, the remaining 40 % of the cyanobacterial community could be described as a cocktail of cHAB genera. Approximately 30 % of all cyanobacterial ASVs were annotated as *Raphidiopsis, Synechocystis, Microcystis, Limnolyngbya* or *Planktothrix* spp., with metagenomic analyses suggesting *Microcystis* spp. populations may have been underrepresented by metabarcoding. Yet, cHAB genera were not blooming at the time of sampling, raising the question which factors constrain cyanobacterial competition within Lake Naivasha.

We found evidence that methane may shape competition dynamics in Lake Naivasha by selecting for non-toxic picoplankton (*Synechococcus* and *Cyanobium* spp.). Indeed, methanogenesis is more substantial in tropical lakes compared to temperate due to elevated productivity rates and benthic sources (Bastviken et al., 2004; Morana et al., 2020). Lake Naivasha possesses a variety of methane sources (most notably papyrus-dominated wetlands) that have been attributed to elevated methane levels (Bastviken et al., 2004; Jones, 2000). Yet, while most studies consider methanogenic bacteria to be major biogenic methane sources - cyanobacteria represent a grossly neglected producer of methane within freshwater systems (Bižić et al., 2020; Kallistova et al., 2023; Yan et al., 2019). A recent study reported "extraordinarily" high concentrations (156 μmol • L^{-1}) of methane in the surface waters of adjacent Lake Sonachi and demonstrated it originated from cyanobacteria including Synechococcales, *Cyanobium* and *Microcystaea* (Fazi et al., 2021). Yet, while some cyanobacteria are known to be a source of methane - the metabolic processes behind this phenomenon remain unclear. Here, we suggest methylphosphonate catabolism by *phn* genes serves as a potential source of methane in Lake Naivasha. We noted a large number of *phn* genes were significantly more abundant in Lake Naivasha's *Synechococcus* and *Cyanobium* spp. genera. Prior studies

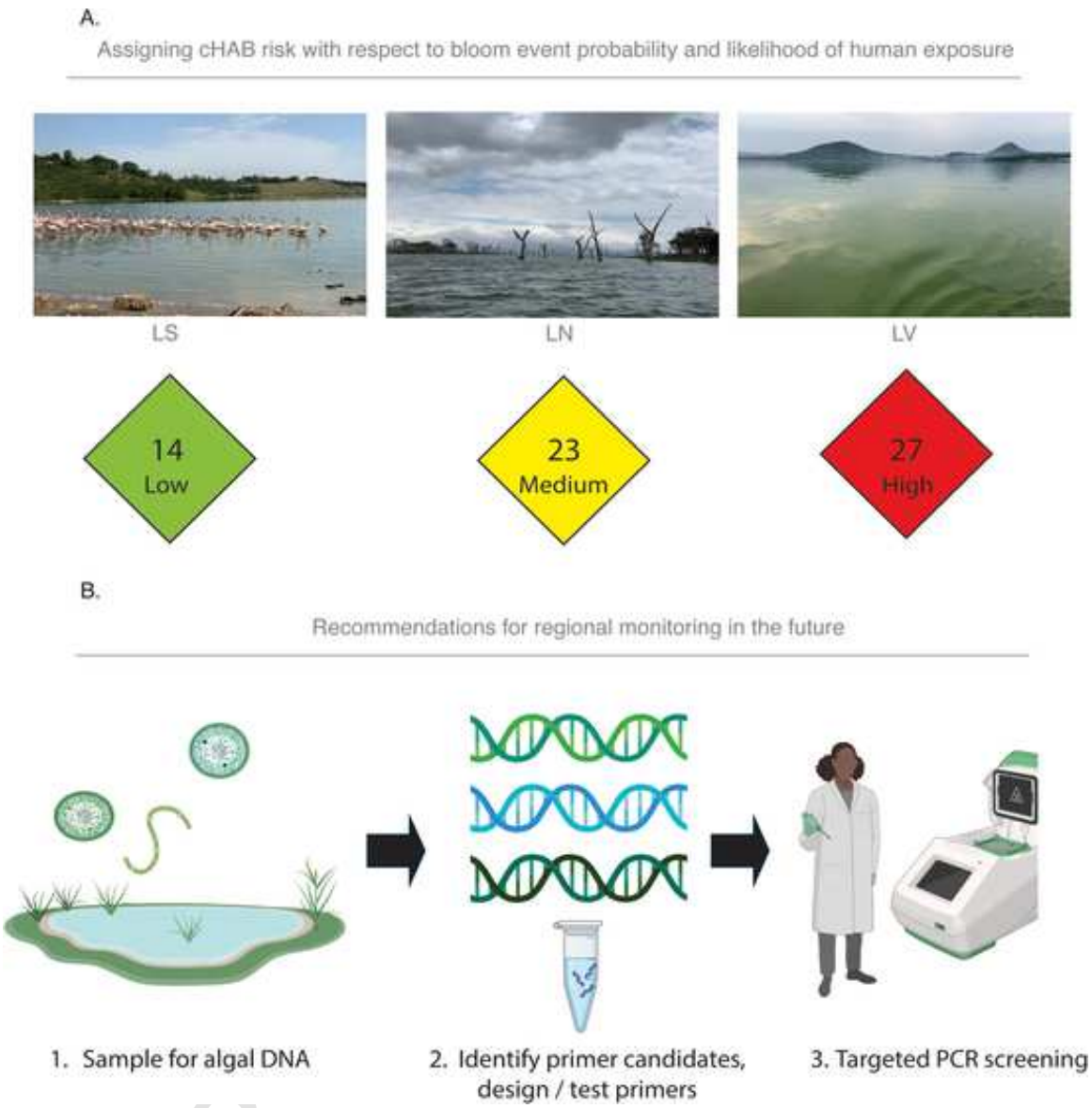

Fig. 8. Proposed framework for estimating cHAB risk and recommendations for future monitoring. (A) Risk scores and designations calculated for Lake Simbi (LS), Lake Naivasha (LN) and Lake Victoria (LV). Scores were based on numerical designations for the ten risk metrics used in this study: Trophic status, water temperature, residence time, wet / dry season, watershed land use, riverine inputs, mean lake depth, basin population, potable water source and recreational use. Here, low risk (green) is a score of <15, moderate risk (yellow) is a score of 16-24, and high risk (red) is a score of >25 (Supplemental Table 25). (B) Schematic of recommendations for local monitoring. Here, we sampled the DNA of three Kenyan lakes of interest (1) and identified potential primer candidates (known, harmful BGC nucleic acid sequences, 2) for future targeted PCR screening (3). Images in panel B created with Biorender.

have shown cyanobacteria that possess *phn* genes grow rapidly in the presence of methylphosphonate (using it as a phosphate source while producing methane) (Teikari et al., 2018; Zhao et al., 2022). In our study, we suggest *Synechococcus* and *Cyanobium* spp. are utilizing *phn* genes to evade phosphorus limitation and exploit Naivasha's nutrient-limited ecological niche. Yet, we only detected putative, *Synechococcus* spp. methylphosphonate transporters in the present study. Hence, further research is required to deduce if Naivasha's *Synechococcus* genera possess *phnJ* to fully metabolize methylphosphonate or if the *phnCDE* genes serve as surrogate inorganic phosphate transporters in the absence of *phnJ* (Martín and Liras, 2021; Saxton et al., 2011; Shah et al., 2023; Zhang et al., 2024). In contrast, *Raphidiopsis* spp. was the only cyanobacteria that possessed transporters (*phnC)* and genes necessary for methylphosphonate catabolism (*phnIJKL)* in the present study. In support, a prior study suggested tropical *Raphidiopsis spp.* was able to utilize phosphonates due to its possession of the complete phosphonate metabolism pathway (Zhi-Jiang et al., 2022). Hence, this suggests *Raphidiopsis* spp. may serve as a formidable competitor to exploit the ecological niche of Lake Naivasha during prolonged phosphorus-limitation. In turn, it remains unclear if *Synechococcus* and *Cyanobium* spp. are metabolizing phosphonates *via* the *phn* gene cluster after import or by another means. Broadly, we suggest methylphosphonate catabolism serves as a cyanobacterial-derived methane source in Lake Naivasha.

While it is widely accepted that cyanobacteria can serve as a methane source (Bižić et al., 2020; Mao et al., 2024; Yan et al., 2019), few have investigated if they function as a methane sink. In the present study, we found *Synechococcus* spp. (and to a lesser extent *Cyanobium* spp.) possessed a wide array of genes involved in the second step of methane metabolism (formaldehyde assimilation through the serine

and ribulose monophosphate pathways) with these genes significantly enriched in Lake Naivasha while they were largely absent in the other lakes. In addition, 2-oxocarboxylic acid metabolism was also enriched in Naivasha's picoplankton, with 2-oxocarboxylic acid chain extension serving as another pathway in methane metabolism. Cumulatively, this suggests picoplankton within Naivasha are metabolizing the product of methane (formaldehyde) and subsequently incorporating the product into biomass. More broadly, this data indicates methanotrophy may uniquely shape cyanobacterial communities in Lake Naivasha while also indicating a potential role of cyanobacteria as a methane sink. Yet, despite this abundance of methanotrophic genes - they did not possess methane transporters or methane monooxygenases -suggesting *Synechococcus* and *Cyanobium* spp. may rely on methane-oxidizing bacteria within the water column or phycosphere to perform initial import and metabolism steps. Indeed, Cerbin et al. (2022) demonstrated methane-oxidizing bacteria could provide alternative carbon sources to support cyanobacterial growth in freshwater systems. In further support, Fazi et al. (2021) noted a bloom of cyanobacteria co-occurred with methanogens - deducing these bacteria were positioned on suspended aggregates which facilitated metabolic interactions between them. In turn, methanogenic archaea and methanotrophic bacteria have been found to co-occur with *Microcystis* spp. aggregates on a global scale (Li et al., 2021). Cumulatively, our data suggest that methane metabolism may shape Lake Naivasha's cyanobacterial community, with *Synechococcus* spp. possessing a notable amount of methanotrophic genes which implies a relationship with methane-oxidizing bacteria and/or the use of methane derivatives. As a result, we hypothesize non-toxic picoplankton in Lake Naivasha may represent a methane sink, with methanotrophy shaping community competition dynamics.

Despite relatively minor abundances of *Microcystis* and *Planktothrix* spp. within Lake Naivasha, BGCs for a variety of *Microcystis* secondary metabolites and *Planktothrix*-synthesized anabaenopeptin were present - though the *mcy* BGC was not detected. While their toxicity is lesser compared to microcystins, anabaenopeptins are toxic towards zooplankton, crustaceans and *C. elegans* due to enzyme inhibition (Lenz et al., 2019) - with synergistic effects observed between anabaenopeptins and other cyanopeptides (Pawlik-Skowrońska and Bownik, 2022). Cumulatively, Lake Naivasha appeared to be nutrient limited at the time of sampling - which may have kept the diversity of toxigenic cHAB taxa at bay and favored phosphonate-laden and methanotrophic *Synechococcus* spp. genera. Yet, while the *mcy* BGC was not detected by our sampling efforts, this does not negate the potential risk of toxigenic cHABs in the future. In turn, additional mining of MEG libraries generated in the present study may reveal novel biosynthetic diversity encoded by these genera.

4.3. Establishing a cHAB risk assessment framework and screening solution

Significant advancements have been made in cHAB prediction and monitoring across freshwater systems such as Lake Erie (NCCOS, 2024). In contrast, there is a lack of cHAB infrastructure in Kenyan lakes and the African continent broadly. This places the anthropogenic and ecological communities that rely on these freshwater resources at substantial risk. Various human health and socioeconomic consequences (including cyanotoxin poisonings) have been documented across coastal cities such as Kisumu and Homa Bay (Obuya et al., 2023; Olokotum et al., 2020; Roegner et al., 2023). In turn, cyanotoxins have been attributed to mass die-offs such as the rapid decline of Lesser Flamingoes (Ndetei and Muhandiki, 2005; Simmons, 2000) and elephant causalities (Lomeo et al., 2024; Veerman et al., 2022, 2024; Zhao et al., 2023). Considering cHAB distribution, duration and frequency are predicted to be exacerbated by climate change - establishing a foundational risk assessment framework and monitoring strategy is imperative. Here, we trial a cHAB risk scoring framework and propose a targeted PCR screening strategy as a preliminary means to address this acute need.

I. *Winam Gulf, Lake Victoria: High Risk.* While classified as hypereutrophic, nitrogen and phosphorus concentrations were low at the time of sampling – suggesting the *Dolichospermum* spp. bloom was approaching maintenance stage as the population self-induced nutrient limitation during the dry season. In turn, low phosphorus concentrations likely kept *Microcystis* spp. at subdominant concentrations (Poste et al., 2013). However, due to the shallow depth of the Gulf and the vast number of rivers transporting nutrients into the water column – it is likely that a replenishment of nutrients (and reseeding of *Microcystis* spp. genera from the sediments) would result in *Microcystis* spp. dominance (Supplemental Table 25) (Table 1). As a result, we hypothesize periods following the wet seasons would be at an elevated risk for *Microcystis* spp. blooms, whereas the dry seasons may favor diazotrophic *Dolichospermum* spp. dominance due to prolonged declines in nutrient replenishment. Cumulatively, we suggest *Microcystis* spp. blooms present a larger threat to ecological and anthropogenic communities as we confirmed this genus possesses the *mcy* BGC, whereas *Dolichospermum* spp. was confirmed to lack toxigenic BGCs at the time of sampling. In turn, due to the various urban centers along the coastline and the direct use of lake water for consumption – the risk for human exposure to cyanotoxins is high (Obuya et al., 2023; Olokotum et al., 2022, 2020). Overall, we recommend the Winam Gulf posits a high cHAB risk and merits routine monitoring.

II. *Lake Naivasha: Moderate Risk.* Though traditionally classified as eutrophic – nutrient profiles were characterized by low nitrogen and phosphorus concentrations at the time of sampling. Yet, unlike the Winam Gulf, there was no evidence of a bloom present and thus cHAB-induced nutrient limitation is unlikely. In turn, it appears the low-nutrient levels may have kept the "cHAB cocktail" at subdominant levels – instead favoring non-toxic picoplankton. Indeed, *Synechococcus* and *Cyanobium* spp. were confirmed to possess *phn* and methanotrophic genes which would enable them to exploit the low-nutrient conditions. However, Lake Naivasha appears to host a harmful cyanobacterial community poised to bloom when environmental conditions become favorable. This lake possessed the largest variety of cHAB genera – including *Microcystis, Raphidiopsis* and *Synechocystis* spp. Hence, once eutrophic conditions return (and this is likely due to rapid, ongoing anthropogenic and agricultural development) the cHAB genera will likely outcompete non-toxic picoplankton and exploit this niche. In turn, because Lake Naivasha is extremely shallow and possesses the longest residence time of the lakes we studied (Table 1), it is likely to undergo continuous reseeding of cHAB genera from the benthos. In turn, this region is undergoing extreme population growth – with humans relying on Lake Naivasha as a direct potable water source due to the dominance of saline, soda lakes in the Rift Vally. As a result, the risk of human exposure to cyanotoxins is high – though the *mcy* BGC was not detected at the time of sampling. In total, Lake Naivasha represents a moderate cHAB risk which we predict will exacerbate with eutrophication, eliciting a need for monitoring.

III. *Lake Simbi: Low Risk.* Despite the *Limnospira* spp. bloom detected upon sampling, dissolved nutrient levels were high - resembling a eutrophic water column. Notably, soda lakes tend to exhibit routinely high phosphorus concentrations (Toner and Catling, 2020). Though prior studies have detected microcystins in the water column during *Limnospira* spp. dominance (Ballot et al., 2005), we did not detect any cyanobacterial BGCs in our samples – suggesting *Limnospira* spp. does not possess the microcystin BGC. In further support, subsequent genomic sequencing of a xenic *Limnospira* spp. culture collected at the time we sampled

yielded no microcystin BGCs or cyanotoxin identification (Brown et al., 2024b). Hence, *Limnospira* spp. does not appear to be the source of microcystins - rather it is likely *Microcystis* spp., which were detected at very low abundances during our sampling. Nonetheless, Lake Simbi's eutrophic physiochemical parameters would seemingly serve to favor cHABs – raising questions as to which factors (beyond nutrients) are constraining these genera. We suggest the high salinity, depth, and consistently monomictic, stratified waters of Lake Simbi reduce cHAB risks – though further research is required. In addition, Lake Simbi is in a region that is sparsely populated – with human populations unable to use the water for consumption due to its saline nature. Hence, the risk of human exposure to cHABs is generally low – though we note local communities bathe in the saline waters as they are believed to treat skin diseases. Considering agricultural development is increasing throughout the area and nutrient inputs (and subsequent stratification changes) are likely to be observed during the wet season – the risk of cHABs still merits attention in this system. Yet, according to this study's risk assessment Simbi does not serve as an ideal candidate for routine monitoring.

5. Conclusions and recommendations

This study identified substantial cHAB risks within Lake Victoria's Winam Gulf and Lake Naivasha – with the dominant cyanobacterial communities constrained by varying metabolic pathways and exhibiting differential niche partitioning. Considering the risk assessment scores above, Winam Gulf and Lake Naivasha serve as model candidates for implementing regional cHAB monitoring. In the present study, we employed metabarcoding and metagenomics as two potential methods for local monitoring strategies. However, subsequent discussions with *NSF-IRES* student participants and stakeholders concluded these methods were not logistically feasible at the local scale. In turn, results within this study concluded the presence of known toxigenic cHAB genera (ex. *Dolichospermum* spp. in Winam Gulf and *Microcystis* spp. in Lake Naivasha) do not correlate to the biosynthetic capacity to produce toxins. As a result, we determined targeted PCR screening serves as a logistically feasible and accurate means to identify cHAB toxin potential and mitigate human exposure. In turn, this method can be synergistically paired with total nitrogen and phosphorus concentrations or additional metrics to bolster cHAB risk predictions. We have included the nucleic acid sequences of known cyanotoxin BGCs that pose a risk to human health (Supplemental Table 26), which serve as primer candidates for local agencies to perform target PCR screenings. In conclusion, this work serves as a foundational framework and impetus for future cHAB identification, risk assessment and monitoring within Kenyan lakes.

Data availability

All raw metagenomic sequencing data generated by the 2022 *NSF-IRES Advanced Studies Institute on Water Quality and Harmful Algal Blooms in Lake Victoria, Kenya* is described within Zepernick et al. (2024 in review). Raw metagenomic sequences collected in 2022 and included within this study are available on the NCBI SRA under BioProject PRJNA1110566. The coassembly generated from the 2022 metagenomic libraries (n = 26) is available on Zenodo (Zepernick, 2024). All metadata are available on the Biological and Chemical Oceanography Data Management Office (BCO-DMO) (Bullerjahn et al., 2024).

Funding

We would like to thank the Kenya Marine and Fisheries Research Institute (KMFRI), Kisii University, Technical University of Kenya, and captain Frederick Okello for their collaboration, access, and assistance while on the RV *Uvumbuzi*. This work was supported by an NSF IRES award #1953468 to K.M.O, R.M.L.M., and G.S.B. in addition to funds provided through NIH-NSF awards 1P01ES0289-01 and 1840715 supporting the Great Lakes Center for Fresh Waters and Human Health at Bowling Green State University (G.J.D., S.W.W., R.M.L.M., G.S.B). Further funds for this work were provided by the NSF GRFP DGE-19389092 and UTK SEC Emerging Scholars Fellowship Program (B.N.Z), the NIH F31 1F31ES036421-01 (L.N.H), and the Hans W. Vahlteich Professorship (D.H.S.).

CRediT authorship contribution statement

Brittany N. Zepernick: Writing – original draft, Methodology, Investigation, Formal analysis, Data curation. **Lauren N. Hart:** Writing – review & editing, Conceptualization, Investigation, Methodology. **Emily E. Chase:** Writing – review & editing, Formal analysis, Validation, Data curation. **Kaela E. Natwora:** Writing – review & editing, Data curation, Investigation, Project administration. **Julia A. Obuya:** . **Mark Olokotum:** Writing – review & editing, Methodology, Investigation. **Katelyn A. Houghton:** Writing – review & editing, Data curation, Investigation, Project administration. **E. Anders Kiledal:** Writing – review & editing, Formal analysis, Validation, Data curation. **Cody S. Sheik:** Writing – review & editing, Validation. **David H. Sherman:** Writing – review & editing, Data curation, Supervision, Investigation. **Gregory J. Dick:** Writing – review & editing, Data curation, Investigation, Project administration. **Steven W. Wilhelm:** . **Lewis Sitoki:** Writing – review & editing, Investigation, Conceptualization. **Kefa M. Otiso:** Writing – review & editing, Investigation, Resources. **R. Michael L. McKay:** Writing – review & editing, Investigation, Supervision, Formal analysis. **George S. Bullerjahn:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization. **Dorine Achieng:** . **Katelyn B. Barker:** . **George M. Basweti:** . **Max Beal:** . **Katelyn M. Brown:** . **Aidan Byrne:** . **Ken G. Drouillard:** . **Albert Getabu:** . **Linet I. Kiteresi:** . **Theodore Lawrence:** . **Davide Lomeo:** . **Jared B. Miruka:** . **Samantha Mohney:** . **James Njiru:** . **Pamela Okutoyi:** . **Reuben Omondi:** . **Dennis Otieno:** . **Omondi A. Owino:** . **Winnie Owoko:** . **Bethwell Owuor:** . **Anakalo Shitandi:** . **Jordyn Stoll:** . **Miriam Swaleh:** . **Emma Tebbs:** . **Emily Varga:** . **Ryan S. Wagner:** .

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The data presented within this manuscript were collected as part of an international collaborative effort between North American and African scientists participating in the *NSF-IRES Advanced Studies Institute on Water Quality and Harmful Algal Blooms in Lake Victoria, Kenya* (https://www.agl-acare.org/bgsu-habs). Consortium authors participated in this program throughout 2022 and/or 2023. All 2022 samples were collected by B.N.Z., K.E.N., J.A.O., and M.O. with the 2022 DNA extractions performed by B.N.Z. Algal Torch measurements for estimated cyanobacterial abundance were performed by J.A.O, M.O. and L.S. Metabarcoding analysis was performed by K.A.H. Metagenomic pre-processing was performed by E.A.K. with subsequent metagenomic analysis performed by B.N.Z. and E.E.C. Biosynthetic Gene Cluster (BGC) analysis was performed L.N.H. A portion of the computation for this work was performed on the *University of Tennessee Infrastructure for Scientific Applications and Advanced Computing (ISAAC)* computational resources. We thank the Kenya Marine and Fisheries Research Institute for their expertise, collaboration, and equipment. We also thank the

crew of the RV *Uvumbuzi* for their assistance and insights. We thank Andrew Wood and Robert Hein for their bioinformatic support. Initial manuscript draft was constructed by B.N.Z. All authors contributed to the final manuscript.

Data availability

Data will be made available on request.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.hal.2024.102757.

Appendix

NSF-IRES 2022 Lake Victoria Research Consortium Participants: Dorine Achieng2,6, Katelyn B. Barker2,14, George M. Basweti2,6, Max Beal2,15, Katelyn M. Brown2,14, Aidan Byrne2, 16, Ken G. Drouillard2,13, Albert Getabu2,17, Linet I. Kiteresi2,6, Theodore Lawrence2,18, Davide Lomeo2,16, Jared B. Miruka2,6, Samantha Mohney2,19, James Njiru2,6, Pamela Okutoyi2,11, Reuben Omondi2,17, Dennis Otieno2,6, Omondi A. Owino2,20, Winnie Owoko2,6, Bethwell Owuor2,17, Anakalo Shitandi2,17, Jordyn Stoll2,21, Miriam Swaleh2,22, Emma Tebbs2,16, Emily Varga2,13, Ryan S. Wagner2,14

15University of Wisconsin - Madison, Madison, Wisconsin, USA, 16Department of Geography, King's College, London, United Kingdom,17Kisii University, Kisii, Kenya, 18African Center for Aquatic Research and Education, 19George Mason University, Fairfax, Virginia, USA, 20Department of Medical and Applied Sciences, Sigalagala National Polytechnic Kakamega, Kenya,21Kent State University, Kent, Ohio, USA, 22Technical University of Mombasa, Mombasa, Kenya,

References

Aguilera, A., Gómez, E.B., Kaštovský, J., Echenique, R.O., Salerno, G.L., 2018. The polyphasic analysis of two native Raphidiopsis isolates supports the unification of the genera Raphidiopsis and Cylindrospermopsis (Nostocales, Cyanobacteria). Phycologia 57 (2), 130–146.

Babicki, S., Arndt, D., Marcu, A., Liang, Y., Grant, J.R., Maciejewski, A., Wishart, D.S., 2016. Heatmapper: web-enabled heat mapping for all. Nucleic. Acids Res. 44 (W1), W147–W153.

Ballot, A., Kotut, K., Novelo, E., Krienitz, L., 2009. Changes of phytoplankton communities in Lakes Naivasha and Oloidien, examples of degradation and salinization of lakes in the Kenyan Rift Valley. Hydrobiologia 632, 359–363.

Ballot, A., Krienitz, L., Kotut, K., Wiegand, C., Pflugmacher, S., 2005. Cyanobacteria and cyanobacterial toxins in the alkaline crater lakes Sonachi and Simbi, Kenya. Harmful. Algae 4 (1), 139–150.

Bastviken, D., Cole, J., Pace, M., Tranvik, L., 2004. Methane emissions from lakes: Dependence of lake characteristics, two regional assessments, and a global estimate. Global. Biogeochem. Cycles 18 (4).

Bižić, M., Klintzsch, T., Ionescu, D., Hindiyeh, M., Günthel, M., Muro-Pastor, A.M., Eckert, W., Urich, T., Keppler, F., Grossart, H.-P., 2020. Aquatic and terrestrial cyanobacteria produce methane. Sci. Adv. 6 (3), eaax5343.

Blanc, B., Gerez, C., de Choudens, S.O., 2015. Assembly of Fe/S proteins in bacterial systems: biochemistry of the bacterial ISC system. Biochim. Biophys. Acta (BBA)-Mol. Cell Res. 1853 (6), 1436–1447.

Brown, K.M., Barker, K.B., Wagner, R.S., Ward, C.S., Sitoki, L., Njiru, J., Omondi, R., Achiya, J., Getabu, A., McKay, R.M., Bullerjahn, G.S., Consortium, N.I.L.V.R., 2024a. Bacterial community and cyanotoxin gene distribution of the Winam Gulf, Lake Victoria, Kenya. Environ. Microbiol. Rep. 16 (3), e13297.

Brown, K.M., Ward, C.S., Bullerjahn, G.S., 2024b. Metagenome-assembled genome sequences of two cyanobacterial cultures from Homa Bay County, Kenya. Microbiol. Resourc. Announc. 13 (3), e01205–e01223.

Bullerjahn, G.S., McKay, R.M.L., Zepernick, B.N., Hart, L.N., 2024. Metagenomic Time Series of Winam Gulf, Lake Victoria from 2022-2023 (ASI Lake Victoria project). Biological and Chemical Oceanography Data Management Office (BCO-DMO).

Bushnell, B., 2014. BBMap: a Fast, Accurate, Splice-Aware Aligner. Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States).

Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., Holmes, S.P., 2016. DADA2: High-resolution sample inference from Illumina amplicon data. Nat. Methods 13 (7), 581–583.

Capelli, C., Ballot, A., Cerasino, L., Papini, A., Salmaso, N., 2017. Biogeography of bloom-forming microcystin producing and non-toxigenic populations of Dolichospermum lemmermannii (Cyanobacteria). Harmful. Algae 67, 1–12.

Cerbin, S., Pérez, G., Rybak, M., Wejnerowski, Ł., Konowalczyk, A., Helmsing, N., Naus-Wiezer, S., Meima-Franke, M., Pytlak, Ł., Raaijmakers, C., 2022. Methane-derived carbon as a driver for cyanobacterial growth. Front. Microbiol. 13, 837198.

Chen, S., 2023. Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. Imeta e107.

Chen, S., Zhou, Y., Chen, Y., Gu, J., 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics 34 (17), i884–i890.

Clarke, K., Gorley, R., 2015. Getting started with PRIMER v7. PRIMER-E: Plymouth. Plymouth Marine Laboratory, p. 20.

Dick, G., 2018. Genomic Approaches in Earth and Environmental Sciences. John Wiley & Sons.

Driscoll, C.B., Meyer, K.A., Šulčius, S., Brown, N.M., Dick, G.J., Cao, H., Gasiūnas, G., Timinskas, A., Yin, Y., Landry, Z.C., 2018. A closely-related clade of globally distributed bloom-forming cyanobacteria within the Nostocales. Harmful. Algae 77, 93–107.

Durazzi, F., Sala, C., Castellani, G., Manfreda, G., Remondini, D., De Cesare, A., 2021. Comparison between 16S rRNA and shotgun sequencing data for the taxonomic characterization of the gut microbiota. Sci. Rep. 11 (1), 3030.

Fakhraee, M., Li, J., Katsev, S., 2017. Significant role of organic sulfur in supporting sedimentary sulfate reduction in low-sulfate environments. Geochim. Cosmochim. Acta 213, 502–516.

Fazi, S., Amalfitano, S., Venturi, S., Pacini, N., Vazquez, E., Olaka, L.A., Tassi, F., Crognale, S., Herzsprung, P., Lechtenfeld, O.J., 2021. High concentrations of dissolved biogenic methane associated with cyanobacterial blooms in East African lake surface water. Commun. Biol. 4 (1), 845.

Finlay, B., Curds, C., Bamforth, S., Bafort, J., 1987. Ciliated protozoa and other microorganisms from two African soda lakes (Lake Nakuru and Lake Simbi, Kenya). Archiv für Protistenkunde 133 (1–2), 81–91.

Frazzon, J., Dean, D.R., 2003. Formation of iron–sulfur clusters in bacteria: an emerging field in bioinorganic chemistry. Curr. Opin. Chem. Biol. 7 (2), 166–173.

Gharib, S.M., Khairy, H.M., Abdel-Aal, E.I., El-Sheekh, M., 2023. Phytoplankton Ecosystem Dynamics and Seasonal Changes in African Lakes: the Case of Lake Victoria, Lakes of Africa. Elsevier, pp. 325–350.

Gikuma-Njuru, P., Hecky, R., 2005. Nutrient concentrations in Nyanza Gulf, Lake Victoria, Kenya: light limits algal demand and abundance. Hydrobiologia 534, 131–140.

Gikuma-Njuru, P., Mwirigi, P., Okungu, J., Hecky, R., Abuodha, J., 2006. Spatial-temporal variability of phytoplankton abundance and species composition in Lake Victoria, Kenya: implication for water quality management.

Gikuma-Njuru, P., Guildford, S., Hecky, R., Kling, H., 2013. Strong spatial differentiation of N and P deficiency, primary productivity and community composition between N yanza G ulf and L ake V ictoria (K enya, E ast A frica) and the implications for nutrient management. Freshwat. Biol. 58 (11), 2237–2252.

Gohl, D.M., Vangay, P., Garbe, J., MacLean, A., Hauge, A., Becker, A., Gould, T.J., Clayton, J.B., Johnson, T.J., Hunter, R., 2016. Systematic improvement of amplicon marker gene methods for increased accuracy in microbiome studies. Nat. Biotechnol. 34 (9), 942–949.

Gophen, M., Ochumba, P.B., Kaufman, L.S., 1995. Some aspects of perturbation in the structure and biodiversity of the ecosystem of Lake Victoria (East Africa). Aquat. Living Resour. 8 (1), 27–41.

Gurevich, A., Saveliev, V., Vyahhi, N., Tesler, G., 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29 (8), 1072–1075.

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

Harper, D.M., Morrison, E.H., Macharia, M.M., Mavuti, K.M., Upton, C., 2011. Lake Naivasha, Kenya: ecology, society and future. Freshw. Rev. 4 (2), 89–114.

Harper, D.M., Phillips, G., Chilvers, A., Kitaka, N., Mavuti, K., 1993. Eutrophication prognosis for Lake naivasha, Kenya. Int. Vereinig. Theoret. Angew. Limnol.: Verhandlungen 25 (2), 861–865.

Hart, L.N., Zepernick, B.N., Natwora, K.E., Brown, K.M., Obuya, J.A., Lomeo, D., Barnard, M.A., Okech, E.O., Consortium, N.-I.L.V.R., Kiledal, E.A., Den Uyl, P.A., Olokotum, M., McKay, R.M.L., Droullard, K.G., Sherman, D.H., Sitoki, L., Aichiya, J., Getabu, A., Otiso, K.M., Bullerjahn, G.S., Dick, G.J., 2024. in review. Metagenomics Reveals Spatial Variation in Cyanobacterial Composition, Function, and Biosynthetic Potential in the Winam Gulf, Lake Victoria, Kenya. American Society of Microbiology.

J. Appl. Environ. Microbiol..
Hayombe, P.O., Odede, F.Z., Agong, S.G., Mossberg, L., 2014. Simbi Nyaima: An Interplay of Nature and Mythology in the Lake Victoria Region; Planning and Management for Ecotourism Transformation in Kenya.
Hecky, R., 1993. The eutrophication of lake Victoria. Int. Vereinig. Theoret. Angew. Limnologie: Verhandlungen 25 (1), 39–48.
Hecky, R.E., Mugidde, R., Ramlal, P., Talbot, M., Kling, G.W., 2010. Multiple stressors cause rapid ecosystem change in Lake Victoria. Freshwat. Biol. 55, 19–42.
Hoffman, B.M., Lukoyanov, D., Yang, Z.-Y., Dean, D.R., Seefeldt, L.C., 2014. Mechanism of nitrogen fixation by nitrogenase: the next stage. Chem. Rev. 114 (8), 4041–4062.
Holmer, M., Storkholm, P., 2001. Sulphate reduction and sulphur cycling in lake sediments: a review. Freshwat. Biol. 46 (4), 431–451.
Hubble, D.S., Harper, D.M., 2002. Phytoplankton community structure and succession in the water column of Lake Naivasha, Kenya: a shallow tropical lake. Hydrobiologia 488, 89–98.
Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S.K., Cook, H., Mende, D.R., Letunic, I., Rattei, T., Jensen, L.J., 2019. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic. Acids. Res. 47 (D1), D309–D314.
Hughes, A.C., Orr, M.C., Ma, K., Costello, M.J., Waller, J., Provoost, P., Yang, Q., Zhu, C., Qiao, H., 2021. Sampling biases shape our view of the natural world. Ecography 44 (9), 1259–1269.
Illumina, 2015. Illumina 16S metagenomic sequencing library preparation.
Ivanikova, N.V., Matteson, A.R., Gobler, C.J., Popels, L.C., McKay, M.L., Wilhelm, S.W., Bullerjahn, G.S., 2008. Picoplanktonic cyanobacteria in Lakes Superior and Erie: phylogenies of endemic populations and cultured isolates. Int. Vereinig. Theoret. Angew. Limnologie: Verhandlungen 30 (3), 459–465.
Jacobson, M.R., Cash, V.L., Weiss, M.C., Laird, N.F., Newton, W.E., Dean, D.R., 1989. Biochemical and genetic analysis of the nifUSVWZM cluster from Azotobacter vinelandii. Mol. General Genet. MGG 219, 49–57.
Jones, M.B., 2000. Methane production and emission from papyrus-dominated wetlands. Int. Vereinig. Theoret. Angew. Limnologie: Verhandlungen 27 (3), 1406–1409.
Kallistova, A.Y., Kosyakova, A., Rusanov, I., Kadnikov, V., Beletsky, A., Koval', D., Yusupov, S., Zekker, I., Pimenov, N., 2023. Methane production in a temperate freshwater lake during an intense cyanobacterial bloom. Microbiol. (N. Y.) 92 (5), 638–649.
Kamat, S.S., Williams, H.J., Raushel, F.M., 2011. Intermediates in the transformation of phosphonates to phosphate by bacteria. Nature 480 (7378), 570–573.
Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., Tanabe, M., 2016. KEGG as a reference resource for gene and protein annotation. Nucleic. Acids. Res. 44 (D1), D457–D462.
Kannaujiya, V.K., Sundaram, S., Sinha, R.P., Kannaujiya, V.K., Sundaram, S., Sinha, R.P., 2017. Structural and functional significance of phycobiliproteins. Phycobil.: Recent Dev. Fut. Appl. 21–44.
Kharwar, S., Bhattacharjee, S., Chakraborty, S., Mishra, A.K., 2021. Regulation of sulfur metabolism, homeostasis and adaptive responses to sulfur limitation in cyanobacteria. Biologia (Bratisl.) 76 (10), 2811–2835.
Krienitz, L., Ballot, A., Wiegand, C., Kotut, K., Codd, G.A., Pflugmacher, S., 2002. Cyanotoxin-producing bloom of Anabaena flos-aquae, Anabaena discoidea and Microcystis aeruginosa (Cyanobacteria. Nyanza Gulf of Lake Victoria, Kenya.
Kumaresan, V., Nizam, F., Ravichandran, G., Viswanathan, K., Palanisamy, R., Bhatt, P., Arasu, M.V., Al-Dhabi, N.A., Mala, K., Arockiaraj, J., 2017. Transcriptome changes of blue-green algae, Arthrospira sp. in response to sulfate stress. Algal. Res. 23, 96–103.
Lawrence, T.J., Achieng, A.O., Chavula, G., Haambiya, L.H., Iteba, J., Kayanda, R., Kaunda, E., Ajode, M.Z., Muvundja, F.A., Nakiyende, H., 2023. Future success and ways forward for scientific approaches on the African Great Lakes. J. Great Lakes Res. 102242.
Lehman, J.T., Branstrator, D.K., 1993. Effects of nutrients and grazing on the phytoplankton of Lake Victoria. Int. Vereinig. Theoret. Angew. Limnologie: Verhandlungen 25 (2), 850–855.
Lehman, J.T., Branstrator, D.K., 1994. Nutrient dynamics and turnover rates of phosphate and sulfate in Lake Victoria, East Africa. Limnol. Oceanogr. 39 (2), 227–233.
Lenz, K.A., Miller, T.R., Ma, H., 2019. Anabaenopeptins and cyanopeptolins induce systemic toxicity effects in a model organism the nematode Caenorhabditis elegans. Chemosphere 214, 60–69.
Li, C., Hambright, K.D., Bowen, H.G., Trammell, M.A., Grossart, H.P., Burford, M.A., Hamilton, D.P., Jiang, H., Latour, D., Meyer, E.I., 2021. Global co-occurrence of methanogenic archaea and methanotrophic bacteria in Microcystis aggregates. Environ. Microbiol. 23 (11), 6503–6519.
Li, D., Liu, C.-M., Luo, R., Sadakane, K., Lam, T.-W., 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics 31 (10), 1674–1676.
Li, H., 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics 34 (18), 3094–3100.
Li, X., Dreher, T.W., Li, R., 2016. An overview of diversity, occurrence, genetics and toxin production of bloom-forming *Dolichospermum* (*Anabaena*) species. Harmful. Algae 54, 54–68.
Li, X., Hou, W., Lei, J., Chen, H., Wang, Q., 2023. The unique light-harvesting system of the algal phycobilisome: structure, assembly components, and functions. Int. J. Mol. Sci. 24 (11), 9733.
Liao, Y., Smyth, G.K., Shi, W., 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30 (7), 923–930.
Lomeo, D., Tebbs, E., Babayani, N.D., Chadwick, M., Gondwe, M.J., Jungblut, A., McCulloch, G., Songhurst, A., Morgan, E., Schillereff, D., 2024. Carcass locations implicate cyanobacterial toxicity in a mass elephant mortality event in Botswana.
Lowe-McConnell, R., 1994. The changing ecosystem of Lake Victoria, East Africa.
Mao, Y., Lin, T., Li, H., He, R., Ye, K., Yu, W., He, Q., 2024. Aerobic methane production by phytoplankton as an important methane source of aquatic ecosystems: Reconsidering the global methane budget. Sci. Total Environ. 907, 167864.
Martín, J.F., Liras, P., 2021. Molecular mechanisms of phosphate sensing, transport and signalling in streptomyces and related actinobacteria. Int. J. Mol. Sci. 22 (3), 1129.
Martin, R.M., Wilhelm, S.W., 2020. Phenol-based RNA Extraction from Polycarbonate Filters. protocols.io.
McKay, R.M., Manyala, J.O., Nunan, F., Nyamweya, C., Awuor, F.J., Njiru, J., 2024. Introduction to the special section: Aquatic resources for a sustainable future and blue economy prosperity for the African Great Lakes region. J. Great Lakes Res. 102413.
Melack, J.M., 1979. Photosynthesis and growth of Spirulina platensis (Cyanophyta) in an equatorial lake (Lake Simbi, Kenya) 1. Limnol. Oceanogr. 24 (4), 753–760.
Morana, C., Bouillon, S., Nolla-Ardèvol, V., Roland, F.A., Okello, W., Descy, J.-P., Nankabirwa, A., Nabafu, E., Springael, D., Borges, A.V., 2020. Methane paradox in tropical lakes? Sedimentary fluxes rather than pelagic production in oxic conditions sustain methanotrophy and emissions to the atmosphere. Biogeosciences 17 (20), 5209–5221.
Mugidde, R., Hecky, R.E., Hendzel, L.L., Taylor, W.D., 2003. Pelagic nitrogen fixation in lake Victoria (East Africa). J. Great Lakes Res. 29, 76–88.
NCCOS, N., 2024. Western Lake Erie Harmful Algal Bloom Early Season Projection Bulletin. NOAA National Centers for Coastal Ocean Science.
Ndetei, R., Muhandiki, V.S., 2005. Mortalities of lesser flamingos in Kenyan Rift Valley saline lakes and the implications for sustainable management of the lakes. Lakes Reserv.: Res. Manage. 10 (1), 51–58.
Njagi, D.M., Routh, J., Odhiambo, M., Luo, C., Basapuram, L.G., Olago, D., Klump, V., Stager, C., 2022. A century of human-induced environmental changes and the combined roles of nutrients and land use in Lake Victoria catchment on eutrophication. Sci. Total Environ. 835, 155425.
Njuguna, H., Romero, J.R., Khisa, P., 2006. The effect of turbid inflows into Winam Gulf, Lake Victoria: a 3D modeling study with ELCOM-CAEDYM.
Norici, A., Hell, R., Giordano, M., 2005. Sulfur and primary production in aquatic environments: an ecological perspective. Photosynth. Res. 86, 409–417.
Obuya, J.A., Onyango, H.O., Olokotum, M., Zepernick, B., Natwora, K., Otieno, D., Owoko, W., Achieng, D., Otuo, P.W., Keyombe, J.L., 2023. Socioeconomic consequences of cyanobacterial harmful algal blooms in small-scale fishing communities of Winam Gulf, Lake Victoria. J. Great Lakes Res. 102236.
Ochumba, P.B., 1990. Massive fish kills within the Nyanza Gulf of lake Victoria, Kenya. Hydrobiologia 208, 93–99.
Ochumba, P.B., Kibaara, D.I., 1989. Observations on blue-green algal blooms in the open waters of Lake Victoria, Kenya. Afr. J. Ecol. 27 (1), 23–34.
Oduor, B., 2018. Pollution in Lake Simbi Nyaima sends away flamingos. Daily Nation.
Ojiambo, B.S., Lyons, W.B., 2019. Residence times of major lons in lake naivasha, kenya, and their relationship to lake hydrology. limnology. Climatol. Paleoclimatol. East Afr. Lakes 267.
Olokotum, M., Humbert, J.-F., Quiblier, C., Okello, W., Semyalo, R., Troussellier, M., Marie, B., Baumann, K., Kurmayer, R., Bernard, C., 2022. Characterization of potential threats from cyanobacterial toxins in Lake Victoria embayments and during water treatment. Toxins (Basel) 14 (10), 664.
Olokotum, M., Mitroi, V., Troussellier, M., Semyalo, R., Bernard, C., Montuelle, B., Okello, W., Quiblier, C., Humbert, J.-F., 2020. A review of the socioecological causes and consequences of cyanobacterial blooms in Lake Victoria. Harmful. Algae 96, 101829.
Omara, T., Nagawa, C.B., Kyarimpa, C., Böhmdorfer, S., Rosenau, T., Lugasi, S.O., Matovu, H., Odongo, S., Ssebugere, P., 2023. Lacustrine cyanobacteria, algal blooms and cyanotoxins in East Africa: implications for human and ecological health protection. Phycology 3 (1), 147–167.
Onywere, S., Mironga, J.M., Simiyu, I., 2012. Use of Remote Sensing Data in Evaluating the Extent of Anthropogenic Activities and their Impact on Lake Naivasha. Kenya.
Opiyo, S., 2020. Assessment of the Temporal and Spatial Variability in the Phytoplankton Dynamics of a Tropical Alkaline-saline Lake Simbi Kenya.
Opiyo, S., 2020. Assessment of the temporal and spatial variability in the phytoplankton dynamics of a tropical alkaline-saline Lake Simbi, Kenya. Int. Res. J. Environ. Sci. 9 (4), 14–23.
Opiyo, S., Getabu, A.M., Sitoki, L.M., Shitandi, A., Ogendi, G.M., 2019. Application of the Carlson's trophic state index for the assessment of trophic status of lake Simbi ecosystem, a deep alkaline-saline lake in Kenya. Int. J. Fisheries Aquat. Stud. 7 (4), 327–333.
Otiang'a-Owiti, G.E., Oswe, I.A., 2007. Human impact on lake ecosystems: the case of Lake Naivasha, Kenya. Afr. J. Aquat. Sci. 32 (1), 79–88.
Owino, O.A., Mokua, O.G., James, O.M., Omondi, O.S., Omondi, R., Ombiro, O.J., 2020. Phytoplankton community structure and ecology in Lake Naivasha, Kenya. Int. J. Fisher. Aquat. Stud. 8 (3), 478–483.
Pawlik-Skowrońska, B., Bownik, A., 2022. Synergistic toxicity of some cyanobacterial oligopeptides to physiological activities of Daphnia magna (Crustacea). Toxicon 206, 74–84.
Phillips, A.A., Ulloa, I., Hyde, E., Agnich, J., Sharpnack, L., O'Malley, K.G., Webb, S.M., Schreiner, K.M., Sheik, C.S., Katsev, S., 2023. Organic sulfur from source to sink in low-sulfate Lake Superior. Limnol. Oceanogr. 68 (12), 2716–2732.
Poste, A., Hecky, R., Guildford, S., 2013. Phosphorus enrichment and carbon depletion contribute to high Microcystis biomass and microcystin concentrations in Ugandan lakes. Limnol. Oceanogr. 58 (3), 1075–1088.
Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner, F.O., 2012. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic. Acids. Res. 41 (D1), D590–D596.
Raini, J.A., 2006. The eastern Africa flamingo Lakes: Building partnerships for sustainable resource management.
Renaut, R.W., Owen, R.B., 2023. Lake Naivasha, The Kenya Rift Lakes: Modern and

Ancient: Limnology and Limnogeology of Tropical Lakes in a Continental Rift. Springer, pp. 417–461.
Roegner, A., Sitoki, L., Weirich, C., Corman, J., Owage, D., Umami, M., Odada, E., Miruka, J., Ogari, Z., Smith, W., 2020. Harmful algal blooms threaten the health of peri-urban fisher communities: A case study in Kisumu Bay, Lake Victoria, Kenya. Expos. Health 12, 835–848.
Roegner, A.F., Corman, J.R., Sitoki, L.M., Kwena, Z.A., Ogari, Z., Miruka, J.B., Xiong, A., Weirich, C., Aura, C.M., Miller, T.R., 2023. Impacts of algal blooms and microcystins in fish on small-scale fishers in Winam Gulf, Lake Victoria: implications for health and livelihood.
Romero, J.R., Alexander, R., 2006. Management implications of the physical limnological studies of Rusinga Channel and Winam Gulf in Lake Victoria.
Rouhiainen, L., Vakkilainen, T., Siemer, B.L., Buikema, W., Haselkorn, R., Sivonen, K., 2004. Genes coding for hepatotoxic heptapeptides (microcystins) in the cyanobacterium Anabaena strain 90. Appl. Environ. Microbiol. 70 (2), 686–692.
Saxton, M.A., Morrow, E.A., Bourbonniere, R.A., Wilhelm, S.W., 2011. Glyphosate influence on phytoplankton community structure in Lake Erie. J. Great Lakes Res. 37 (4), 683–690.
Scheren, P., Zanting, H., Lemmens, A., 2000. Estimation of water pollution sources in Lake Victoria, East Africa: application and elaboration of the rapid assessment methodology. J. Environ. Manage. 58 (4), 235–248.
Shah, B.S., Ford, B.A., Varkey, D., Mikolajek, H., Orr, C., Mykhaylyk, V., Owens, R.J., Paulsen, I.T., 2023. Marine picocyanobacterial PhnD1 shows specificity for various phosphorus sources but likely represents a constitutive inorganic phosphate transporter. ISMe J. 17 (7), 1040–1051.
Simmons, R., 2000. Declines and movements of lesser flamingos in Africa. Waterbirds 40–46.
Sitoki, L., Gichuki, J., Ezekiel, C., Wanda, F., Mkumbo, O.C., Marshall, B.E., 2010. The environment of Lake Victoria (East Africa): current status and historical changes. Int. Rev. Hydrobiol. 95 (3), 209–223.
Sitoki, L., Kurmayer, R., Rott, E., 2012. Spatial variation of phytoplankton composition, biovolume, and resulting microcystin concentrations in the Nyanza Gulf (Lake Victoria, Kenya). Hydrobiologia 691, 109–122.
Stark, G.F., Martin, R.M., Smith, L.E., Wei, B., Hellweger, F.L., Bullerjahn, G.S., McKay, R.M.L., Boyer, G.L., Wilhelm, S.W., 2023. Microcystin aids in cold temperature acclimation: Differences between a toxic Microcystis wildtype and non-toxic mutant. Harmful. Algae 129, 102531.
Svirčev, Z., Lalić, D., Bojadžija Savić, G., Tokodi, N., Drobac Backović, D., Chen, L., Meriluoto, J., Codd, G.A., 2019. Global geographical and historical overview of cyanotoxin distribution and cyanobacterial poisonings. Arch. Toxicol. 93, 2429–2481.
Teikari, J.E., Fewer, D.P., Shrestha, R., Hou, S., Leikoski, N., Mäkelä, M., Simojoki, A., Hess, W.R., Sivonen, K., 2018. Strains of the toxic and bloom-forming Nodularia spumigena (cyanobacteria) can degrade methylphosphonate and release methane. ISMe J. 12 (6), 1619–1630.
Tipping, E., Carrick, T., Hurley, M.A., James, J., Lawlor, A., Lofts, S., Rigg, E., Sutcliffe, D., Woof, C., 1998. Reversal of acidification in upland waters of the English Lake District. Environ. Pollut. 103 (2–3), 143–151.
Toner, J.D., Catling, D.C., 2020. A carbonate-rich lake solution to the phosphate problem of the origin of life. Proc. Natl Acad. Sci. 117 (2), 883–888.
Tuite, C., 1979. Population size, distribution and biomass density of the Lesser Flamingo in the Eastern Rift Valley, 1974-76. J. Appl. Ecol. 765–775.
Veerman, J., Kumar, A., Mishra, D.R., 2022. Exceptional landscape-wide cyanobacteria bloom in Okavango Delta, Botswana in 2020 coincided with a mass elephant die-off event. Harmful. Algae 111, 102145.
Veerman, J., Mishra, D.R., Kumar, A., Karidozo, M., 2024. Environmental drivers behind the exceptional increase in cyanobacterial blooms in Okavango Delta, Botswana. Harmful. Algae 102677.
Wagner, G.P., Kin, K., Lynch, V.J., 2012. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. Theory Biosci. 131, 281–285.
Wang, J., Wagner, N.D., Fulton, J.M., Scott, J.T., 2021. Diazotrophs modulate phycobiliproteins and nitrogen stoichiometry differently than other cyanobacteria in response to light and nitrogen availability. Limnol. Oceanogr. 66 (6), 2333–2345.
Wells, M.L., Karlson, B., Wulff, A., Kudela, R., Trick, C., Asnaghi, V., Berdalet, E., Cochlan, W., Davidson, K., De Rijcke, M., 2020. Future HAB science: directions and challenges in a changing climate. Harmful. Algae 91, 101632.
Wilhelm, S., Bullerjahn, G., RLM, M., 2020. The complicated and confusing ecology of *Microcystis* blooms. mBio 11 (3). . e00529-00520.
Wilhelm, S.W., Bullerjahn, G.S., Eldridge, M.L., Rinta-Kanto, J.M., Poorvin, L., Bourbonniere, R.A., 2006. Seasonal hypoxia and the genetic diversity of prokaryote populations in the central basin hypolimnion of Lake Erie: evidence for abundant cyanobacteria and photosynthesis. J. Great Lakes Res. 32 (4), 657–671.
Wishart, 1998. The increasing divide between First and Third Worlds: science, collaboration and conservation of Third World aquatic ecosystems. Freshwat. Biol. 39 (3), 557–567.
Wood, D.E., Lu, J., Langmead, B., 2019. Improved metagenomic analysis with Kraken 2. Genome Biol. 20, 1–13.
Yan, X., Xu, X., Ji, M., Zhang, Z., Wang, M., Wu, S., Wang, G., Zhang, C., Liu, H., 2019. Cyanobacteria blooms: A neglected facilitator of CH4 production in eutrophic lakes. Sci. Total Environ. 651, 466–474.
Yancey, C.E., Mathiesen, O., Dick, G.J., 2023. Transcriptionally active nitrogen fixation and biosynthesis of diverse secondary metabolites by Dolichospermum and Aphanizomenon-like Cyanobacteria in western Lake Erie Microcystis blooms. Harmful. Algae 124, 102408.
Yongo, E., Agembe, S.W., Manyala, J.O., Mutethya, E., 2023. Assessment of the current trophic state and water quality of Lake Naivasha, Kenya using multivariate techniques. Lakes Reserv.: Res. Manage. 28 (1), e12422.
Yu, G., Wang, L.-G., Han, Y., He, Q.-Y., 2012. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS: J. Integr. Biol. 16 (5), 284–287.
Zepernick, B.N., 2024. Lakes Victoria, Simbi, Naivasha Coassembly (Metagenomic).
Zepernick, B.N., Hart, L.N., Natwora, K.E., Brown, K.M., Obuya, J.A., Olokotum, M., Owino, E., Keating, N.G., Lomeo, D., Tebbs, E.J., Consortium., N.-I.-L.V.R., Sheik, C.S., Sherman, D.H., J., D.G., Wilhelm, S.W., Droulliard, K.G., Lawrence, T.J., Getabu, A., Owuor, B., Shitandi, A., Omondi, R., Nijiru, J., Sitoki, L., Otiso, K.M., McKay, R.M.L., Bullerjahn, G.S., 2024 in review. Metagenomic Sequencing of Cyanobacterial-Dominated Lake Victoria - an African Great Lake. ASM MRA.
Zepernick, B.N., McKay, R.M.L., Martin, R.M., Bullerjahn, G.S., Paerl, H.W., Wilhelm, S.W., 2024. A tale of two blooms: do ecological paradigms for algal bloom success and succession require revisiting? J. Great Lakes Res. 102336.
Zepernick, B.N., Truchon, A.R., Gann, E.R., Wilhelm, S.W., 2022. Draft genome sequence of the freshwater diatom *Fragilaria crotonensis* SAG 28.96. Microbiol. Resour. Announc. 11 (9). . e00289-00222.
Zepernick, B.N., Wilhelm, S.W., Bullerjahn, G.S., Paerl, H.W., 2023. Climate change and the aquatic continuum: A cyanobacterial comeback story. Environ. Microbiol. Rep. 15 (1), 3–12.
Zhang, Q., Jia, L., Chen, Y., Yan, H., Chen, Q., Zhang, J., Sun, H., 2024. Molecular mechanisms of the cyanobacterial response to different phosphorus sources. Sustainability 16 (13), 5642.
Zhao, L., Lin, L.-Z., Chen, M.-Y., Teng, W.-K., Zheng, L.-L., Peng, L., Lv, J., Brand, J.J., Hu, C.-X., Han, B.-P., 2022. The widespread capability of methylphosphonate utilization in filamentous cyanobacteria and its ecological significance. Water. Res. 217, 118385.
Zhao, X., Liu, Y., Guo, Y.-M., Xu, C., Chen, L., Codd, G.A., Chen, J., Wang, Y., Wang, P.-Z., Yang, L.-W., 2023. Meta-analysis reveals cyanotoxins risk across African inland waters. J. Hazard. Mater. 451, 131160.
Zhi-Jiang, C., Zi-Xi, R., Nan, C., Li-Juan, X., Liang, P., Bo-Ping, H., La-Mei, L., 2022. Whole-genome sequencing and phosphorus uptake and transport pathway comparative analysis of Cylindrospermopsis raciborskii N8. Acta Hydrobiol. Sinica 46 (8), 1130–1141.
Zhu, W., Lomsadze, A., Borodovsky, M., 2010. Ab initio gene identification in metagenomic sequences. Nucleic Acids. Res. 38 (12). . e132-e132.
Zilliges, Y., Kehr, J.-C., Meissner, S., Ishida, K., Mikkat, S., Hagemann, M., Kaplan, A., Boerner, T., Dittmann, E., 2011. The cyanobacterial hepatotoxin microcystin binds to proteins and increases the fitness of Microcystis under oxidative stress conditions. PLoS One 6 (3), e17615.