Environmental Microbiology

ENVIRONMENTAL MICROBIOLOGY

RESEARCH ARTICLE OPEN ACCESS

Genetic Diversity and Anatoxin Profiles of Freshwater Benthic Cyanobacteria From Nova Scotia (Canada)

Cecilio Valadez-Cano[1] | Nicolas Tromas[2] | Adrian Reyes-Prieto[1] | Lindsay Johnston[3] | Yannan Huang[3] | Hannah Morris[3] | Lydia Zamlynny[3,4] | Daniel G. Beach[3,4] | Rob C. Jamieson[3] | Janice Lawrence[1]

[1]Department of Biology, University of New Brunswick, Fredericton, New Brunswick, Canada | [2]INRAE, CARRTEL, Thonon-les-Bains, France | [3]Centre for Water Resources Studies, Department of Civil & Resource Engineering, Dalhousie University, Halifax, Nova Scotia, Canada | [4]Biotoxin Metrology, National Research Council Canada, Halifax, Nova Scotia, Canada

Correspondence: Janice Lawrence (jlawrenc@unb.ca)

Received: 23 November 2024 | **Revised:** 30 January 2025 | **Accepted:** 1 February 2025

Funding: This work was supported by Natural Sciences and Engineering Research Council of Canada, ALLRP 576532—22.

Keywords: anatoxin-a | benthic cyanobacteria | dihydroanatoxin-a | dihydrohomoanatoxin-a | homoanatoxin-a | metagenome | *Microcoleus*

ABSTRACT

Some mat-forming cyanobacteria produce harmful cyanotoxins, yet benthic species remain understudied compared to planktonic counterparts. This study assesses the diversity, distribution and toxin production of mat-forming cyanobacteria across lentic and lotic systems in Nova Scotia, Canada. We documented greater cyanobacterial species richness in lentic environments, with six dominant species distributed into two major *Microcoleus* clades, five of which represent putative novel taxa. Two *Microcoleus* species with the genetic repertoire to produce anatoxins were prevalent. One has been previously reported in Canada, while the second represents a novel species found exclusively in an environment impacted by discharge from a water treatment plant. We observed variability in the gene clusters responsible for the biosynthesis of anatoxin-a and associated analogues (ATXs), including the discovery of a novel *anaG* variant with a ~1.7 kb insertion in a *Microcoleus* strain dominating homoanatoxin-producing mats. This extended *anaG*, encoding a polyketide synthase with an additional methyltransferase domain, coexists with shorter variants, leading to the production of a mixture of ATXs. These findings highlight the genetic diversity of benthic cyanobacteria in freshwater environments, with *Microcoleus* as the primary contributor to the production of ATXs in both lentic and lotic systems, underscoring their potential to produce harmful toxins.

1 | Introduction

The proliferation of benthic mats producing cyanotoxins, which are associated with animal poisonings in freshwater aquatic systems, has been documented for more than 30 years (Edwards et al. 1992; Gunn et al. 1992). Among the cyanotoxins produced by benthic cyanobacteria, anatoxin-a (ATX) and its analogues (hereafter collectively referred to as ATXs) are potent neurotoxins that act as agonists of the nicotinic acetylcholine receptor, causing rapid death when ingested, even in low doses (World Health Organization 2020). Production of ATXs has been reported in various cyanobacterial taxa, including *Microcoleus* species, inhabiting both lentic and lotic systems (Faassen et al. 2012; Wood et al. 2012; Quiblier et al. 2013). Ingestion of mats dominated by toxigenic *Microcoleus* has been implicated in an increasing number of domestic animal and wildlife deaths (Gugger et al. 2005; Faassen et al. 2012; Wood et al. 2017; Bouma-Gregson et al. 2019; McCarron et al. 2023; Johnston et al. 2024). These incidents associated with *Microcoleus* have prompted growing interest in exploring the genetic diversity, geographic distribution, ecological roles, and toxin-producing potential of this genus.

In Canada, the first reports of canine poisonings associated with ATXs produced by benthic *Microcoleus* came from the

Wolastoq shores in the Fredericton area, New Brunswick (McCarron et al. 2023). Subsequent analyses revealed that toxigenic *Microcoleus* species were prevalent at various sites along the Wolastoq benthos (Beach et al. 2022), with their proliferation showing a clear temporal dynamic influenced by biotic factors, such as phage infection (Valadez-Cano, Reyes-Prieto, and Lawrence 2023). Similarly, in the neighbouring province of Nova Scotia, dog deaths attributed to ATXs were reported in 2021 and 2023 along the shores of Shubenacadie-Grand Lake and Long Lake, respectively (Crosby 2023; Johnston et al. 2024).

Recent studies have shown that *Microcoleus*-dominated mats and derived isolates can produce various ATXs. For example, *Microcoleus anatoxicus* from diverse geographic regions has been found to predominantly synthesise dihydroanatoxin-a (H_2ATX) (Conklin et al. 2020; Junier et al. 2024; Stancheva et al. 2024). In Nova Scotia, benthic mats dominated by *Microcoleus* along a stream-lake continuum in Shubenacadie-Grand Lake and Long Lake revealed substantial production of homoanatoxin-a (hATX), alongside other analogues (Johnston et al. 2024). *Microcoleus*-dominated mats from the Wolastoq were found to produce ATX and H_2ATX, with ATX primarily detected in isolates derived from these mats (Valadez-Cano, Reyes-Prieto, Beach et al. 2023). Additionally, significant production of different ATXs, including hATX and dihydrohomoanatoxin-a (H_2hATX), has been observed in mats and isolates from New Zealand (Heath et al. 2011; Wood et al. 2012; Kelly et al. 2018).

The production of ATXs by *Microcoleus* has been linked to the presence of a cluster of genes, the *ana* cluster, identified in other cyanobacteria (Méjean et al. 2014). This includes three multidomain polyketide synthases (PKSs; *anaE*, *anaF* and *anaG*) involved in the catalysis of ATXs from proline (Méjean et al. 2014). Despite the variability in ATXs observed, the *ana* clusters reported in *Microcoleus* to date are highly conserved in both protein-coding content and arrangement (Bouma-Gregson et al. 2019; Conklin et al. 2020; Tee et al. 2021; Valadez-Cano, Reyes-Prieto, Beach et al. 2023; Junier et al. 2024). These highly similar *ana* clusters do not fully explain the observed differences in the production of ATXs, which have been associated with specific variations in the *ana* clusters of other cyanobacteria (Méjean et al. 2014). For instance, in *Kamptonema* sp. PCC 6506 (formerly *Oscillatoria* sp. PCC 6506), hATX is derived from a precursor that is methylated by the PKS *anaG* (Méjean et al. 2009). Despite recurrent reports of hATX in *Microcoleus*-dominated mats and isolates, the *anaG* gene described in *Microcoleus* is shorter than its homologue in *Kamptonema* sp. PCC 6506 and does not encode the methyltransferase domain (Bouma-Gregson et al. 2019; Conklin et al. 2020; Tee et al. 2021; Valadez-Cano, Reyes-Prieto, Beach et al. 2023; Junier et al. 2024). Similarly, biosynthesis of H_2ATX in *Cylindrospermum stagnale* PCC 7417 is proposed to involve the reduction of either ATX or its precursor through an F_{420}-dependent oxidoreductase domain encoded by the *anaK* gene (Méjean et al. 2016). With the exception of some *M. anatoxicus* strains reported in California, all *Microcoleus ana* clusters reported so far contain at least one copy of the *anaK* gene, though some strains produce low concentrations of H_2ATX (Bouma-Gregson et al. 2019; Conklin et al. 2020; Tee et al. 2021; Valadez-Cano, Reyes-Prieto, Beach et al. 2023; Junier et al. 2024).

In this study, we investigated benthic mats from freshwater environments across various ecoregions of Nova Scotia to identify ATXs and their producers. Samples were collected from lentic and lotic systems with differing trophic states, and from various substrates during 2022 and 2023. We quantified the ATXs and genetically characterised the cyanobacteria associated with these mats. Additionally, we examined the variation in the composition and synteny of the *ana* clusters to explore differences potentially associated with the production of anatoxin-a and its analogues.

2 | Methods

2.1 | Sample Collection and DNA Extraction

Although Nova Scotia is one of the smallest provinces in Canada (~55,000 km^2), its varied physiographic and climatic characteristics result in a diversity of freshwater environments (Johnston et al. 2022). During the summer seasons of 2022 and 2023, 20 periphyton samples suspected to be dominated by cyanobacteria were collected from nine freshwater systems across the province. These systems encompassed both lentic environments (Fish Lake, Grand Lake and Oathill Lake) and lotic environments (Byers Brook, French River, Great Village River, Indian Brook, Little Sackville River and North River) (Figure S1). Sampling was conducted across three distinct ecoregions (Nova Scotia Uplands, Eastern and Cape Breton Highlands) (Neily et al. 2017), spanning a gradient of trophic states (Table S1). Benthic cyanobacterial mats were visually identified in the field, and multiple subsamples were collected using a gloved hand and composited together into a clean, sterile plastic jar. The type of substrate on which the mats were growing was noted during sample collection (Table S1). Samples were transported to the laboratory in a chilled cooler. In the laboratory, mat samples were microscopically examined (400× magnification with a Nikon Eclipse 80i microscope equipped with a Nikon DS-Fil camera) to investigate the presence and abundance of cyanobacteria, and if the detected specimens morphologically resembled members of the genera *Oscillatoria*, *Phormidium*, or *Microcoleus* (Figure S2). If the target taxa appeared dominant, mat samples were homogenised using 15 mL sterile tissue grinders (VWR, Mississauga, Ontario, Canada). Then, subsamples of 0.25 g of each homogenate were subjected to DNA extraction using the protocol of the Qiagen DNeasy PowerSoil Pro kits (Qiagen Inc., Toronto, Ontario, Canada).

2.2 | Quantification of Anatoxin-a and Analogues

To assay total intracellular ATXs, subsamples of homogenised mat (1 g) were frozen and thawed prior to toxin extraction to lyse cyanobacterial cells. Then toxins were extracted with 1:1 g mL^{-1} methanol containing 0.1% formic acid, vortex mixed for 1 min, centrifuged for 10 min at 21000 g and filtered to 0.45 μm using 5 mL PVDF centrifugal filters (Millipore, Bedford, MA, USA). ATXs were measured by

liquid chromatography–high resolution tandem mass spectrometry (LC–HRMS/MS) using a method reported recently (Beach et al. 2023). Briefly, aliquots of sample extract (90 μL) were spiked with 10 μL of 600 ng mL^{-1} $^{13}C_4$-ATX standard (Eurofins Abraxis, Warminster, PA, USA). An Agilent 1290 Infinity II liquid chromatography system (Agilent, Santa Clara, CA, USA) was coupled to a Q Exactive HF mass spectrometer with a HESI-II electrospray source (ThermoFisher Scientific, Waltham, MA, USA). The LC separation used an Acquity 1.8 μm HSS T3 C18 column (150 × 2.1 mm; Waters, Milford, MA, USA) and a linear gradient from 2% to 11% acetonitrile (B) in water (A), both containing 0.1% formic acid, over 25 min, at a flow rate of 600 μL min^{-1}. This was followed by a 10-min column flush with 95% B at 600 μL min^{-1} and a 10-min re-equilibration at initial gradient conditions. An injection volume of 1 μL was used for all samples. Acquisition was performed in data-dependent mode, capturing full scans over an *m/z* range of 140–500 at a resolution of 60,000. Up to five product ion spectra were generated per cycle at a resolution of 15,000, with a collision energy of 20 eV. Certified reference material calibration solutions for ATX (CRM-ATX) and hATX (CRM-hATX) were obtained from the National Research Council of Canada (Halifax, NS). A standard for H_2ATX was obtained from Gold Standard Diagnostics (Warminster, PA). Quantitation of ATX was carried out as described previously by isotope dilution for ATX and external calibration for hATX and H_2ATX (Beach et al. 2022), with the addition of H_2hATX semi-quantitated using the hATX standard. Total ATXs were measured as the sum of ATX, H_2ATX, hATX and H_2hATX. ATXs were measured in units of mg kg^{-1} wet weight of drained mat material, as reported previously (Beach et al. 2022, 2023), which was considered appropriate due to the very high biological variability in the concentration of ATXs within the mats (Beach et al. 2022, 2023).

2.3 | Shotgun Metagenomic Sequencing

Total DNA purified from each benthic mat sample was shotgun sequenced using an Illumina NextSeq 550 platform, generating 150 bp paired-end reads. The library preparation and DNA sequencing procedures were carried out at Dalhousie University's Integrated Microbiology Resource in Halifax, NS, following standard protocols (Comeau and Filloramo 2023).

2.4 | Read Processing, Assembly and Recovery of Metagenome-Assembled Genomes

Illumina adapters and low-quality reads were removed with *fastp v0.20.1* (Chen et al. 2018) using default settings. Cleaned reads were assembled into contigs and scaffolds using *metaSPAdes* (*SPAdes v3.12.0*) (Bankevich et al. 2012), using *k-mer* options -k21,33,55,77. Metagenome-assembled genomes (MAGs) were recovered from each assembly using the *variational autoencoders for metagenomic binning pipeline* (*VAMB v3.0.2*) (Nissen et al. 2021), *MetaBAT2 v2.15-6* (Kang et al. 2015) and *MaxBin v2.2.7* (Wu et al. 2016). Then, *DasTool v1.1.4* (Sieber et al. 2018) was used to recover the highest-quality non-redundant MAGs from each assembly. MAG quality was estimated with *CheckM v1.1.3* (Parks et al. 2015).

2.5 | Phylogenomic Analysis and Cyanobacterial Species Cluster Definition

Taxonomic classification of MAGs was performed with *GTDB-Tk v2.4.0* (Chaumeil et al. 2022) based on the Genome Database Taxonomy (GTDB) r220 (Parks et al. 2022). We used *Prodigal v2.6.3* (Hyatt et al. 2010) to predict Open Reading Frames (ORFs) from cyanobacterial MAGs (cMAGs) with an estimated completeness > 70% and contamination < 5% recovered from Nova Scotia (Table S2), as well as from previously documented cyanobacterial genomes. To elucidate the phylogenetic relationships of the recovered cMAGs from Nova Scotia, a comprehensive analysis was conducted considering reported genomes of *Microcoleus* species and other cyanobacteria with or without the *ana* cluster (Table S3, DataSet1). *Orthofinder v2.5.5* (Emms and Kelly 2015, 2019) was used to construct a protein multiple sequence alignment (MSA). Then, a maximum-likelihood (ML) tree was inferred from the previously generated MSA with *IQ-TREE v2.3.2* (Nguyen et al. 2015) considering the Q.plant + F + I + R6 model, which was identified as the fittest amino acid substitution model by *ModelFinder* (Kalyaanamoorthy et al. 2017). Node support was evaluated with 2000 ultrafast bootstrap replicates.

A genome-sequence-based species delimitation analysis was performed on cMAGs recovered from Nova Scotia mat samples, and other *Microcoleus* and cyanobacteria reported from the USA, New Zealand, and Canada (Table S3, DataSet2) (Bouma-Gregson et al. 2019; Tee et al. 2021; Valadez-Cano, Reyes-Prieto, Beach et al. 2023; Junier et al. 2024). The average nucleotide identity (ANI) between genomes was estimated with *FastANI v1.33* (Jain et al. 2018). The digital DNA–DNA hybridisation (dDDH) index between pairs of genomes was estimated with the *Genome-to-Genome Distance Calculator* web-based DSMZ service (https://ggdc.dsmz.de/) (Meier-Kolthoff et al. 2013). Cut-offs of 95% ANI and 70% for dDDH were used to delineate bacterial species, following previously suggested criteria (Richter and Rosselló-Móra 2009; Meier-Kolthoff et al. 2013).

2.6 | Prediction of the *ana* Cluster in Cyanobacterial Genomes

The online bacterial version of *antiSMASH* (Blin et al. 2021) was used to predict the presence of the *ana* cluster in the recovered cMAGs. Multiple alignments and visualisations of the different *ana* clusters were generated with *DiGAlign* (Nishimura et al. 2024).

Sequencing coverage (i.e., read mapping) of the *ana* cluster from each mat sample was assessed with *Bowtie2 v2.3.5.1* (Langmead and Salzberg 2012). Read mapping visualisation of samples of interest was performed with the *anvi'o v8* tool (Eren et al. 2021).

2.7 | Relative Abundance and Diversity of Cyanobacterial Species Clusters

The MIDAS pipeline (Nayfach et al. 2016) was used to estimate the relative abundance of cyanobacterial species clusters

identified in our genome-based species delimitation analysis. First, a custom MIDAS database was built for each predefined cyanobacterial species cluster. Specifically, for each species cluster, a non-redundant pangenome was created by clustering genes at 99% sequence identity and identifying a centroid gene sequence from each cluster using *VSEARCH* (Rognes et al. 2016). Then, a database for marker genes was built based on 15 single-copy gene families. Finally, with *MIDAS*, we mapped the cleaned reads from each sequenced mat sample to the custom database to estimate the relative abundance of each cyanobacterial species cluster in all the mat samples.

The Shannon diversity index was calculated with the *estimate_richness* function of the phyloseq R package *v1.44.0* (McMurdie and Holmes 2013) to compare the cyanobacterial diversity across environmental variables (Table S1). We then used the square root of Jensen–Shannon divergence (JSD) with the *distance* function from the phyloseq vegan R package *v2.6.4* (Oksanen et al. 2022) to investigate the spatiotemporal variations in the cyanobacterial community composition. Differences in these variables between groups were also tested using permutational multivariate analysis of variance (PERMANOVA) (Anderson 2001) with the *adonis2* function. We also tested for dispersion in the data by performing an analysis of multivariate homogeneity (PERMDISP) (Anderson 2006) with the permuted *betadisper* function. PERMANOVA and PERMDISP were performed in R. The non-parametric Analysis of Similarity (ANOSIM) (Clarke 1993) with the *anosim* function was implemented to validate the observed groups. To identify cyanobacterial taxa specific to lentic or lotic systems, the cyanobacterial abundance table was initially transformed using the Hellinger transformation (Rao 1995) via the *transform* function from the microbiome R package; then, an indicator species analysis with the *multipatt* function from the *indicspecies v1.7.14* R package (Cáceres and Legendre 2009) was used. Finally, a redundancy analysis was performed to identify cyanobacterial taxa associated with the specific environmental variables and concentrations of ATXs (Rao 1964).

3 | Results

3.1 | Anatoxin-a and Analogue Quantification in Mat Samples From Nova Scotia

With exception of mat NSLS-A, ATXs were detected in the remaining 19 mat samples (Table S4), with total concentrations varying both between and within different aquatic systems. Concentrations exceeded 1 mg kg^{-1} in 55% of the samples, with these high toxin samples having been collected from Oathill and Grand lakes, as well as from Indian Brook, Great Village and French rivers (Figure 1A).

ATXs were detected in varying proportions across samples, with some aquatic systems producing specific analogues (Figure 1B). For instance, only ATX was detected in mat samples from Byers Brook, whereas Grand Lake mats primarily contained H_2ATX (Figure 1B). Although ATX and H_2ATX were the predominant analogues in most samples, hATX and H_2hATX constituted over 98% of the total ATXs in samples NSOL-A and NSOL-B collected from Oathill Lake (Figure 1B). Three out of five mats collected from French River contained ATX and H_2ATX, while mat NSFL-A from Fish Lake predominantly contained hATX, with H_2ATX comprising 17% of the total ATXs. Sample NSIB-A from Indian Brook contained a total of 4.4 mg kg^{-1} of ATXs, with hATX and H_2hATX accounting for 16% of the total ATXs, although ATX and H_2ATX were the primary analogues detected (Figure 1B).

3.2 | Diversity of Mat-Forming Cyanobacteria

A total of 25 cMAGs with completeness exceeding 75% and contamination below 5% were recovered (Table S2). Between 1 and 3 cMAGs were obtained from each mat, except for NSGL-C (Grand Lake), where no cyanobacterial genomes were assembled (Table S2). Sixteen cMAGs were classified within the *Microcoleus* genus (Microcoleaceae family), while eight cMAGs were assigned to the Coleofasciculaceae family, with one of these further classified as belonging to the *Limnofasciculus* genus (Table S2). Additionally, a single cMAG was identified as belonging to the *Pseudanabaena* genus (Pseudanabaenaceae family) (Table S2). In agreement with the taxonomic classification of the cMAG, our maximum-likelihood (ML) tree resolved 16 cMAGs from Nova Scotia into two major clades within the *Microcoleus* genus (100% bootstrap; BS) (Figure 2). The remaining nine cMAGs retrieved from Nova Scotia were affiliated with other cyanobacterial groups and were all non-toxigenic (Figure 2).

The first major *Microcoleus* clade (MC1; 100% BS) unites 7 non-toxigenic cMAGs from five distinct aquatic systems in Nova Scotia with non-toxigenic *Microcoleus* genomes reported from the United States and New Zealand (Bouma-Gregson et al. 2019; Tee et al. 2021) (Figure 2). Additionally, this clade includes genomes from recently described *Microcoleus* clades (M1-M12) recovered from diverse soil, moss, marine, and puddle habitats (Stanojković et al. 2024). The second major *Microcoleus* clade (MC2; 100% BS) comprises *Microcoleus* genomes recovered only from rivers or lakes (Figure 2). Within this clade, nine cMAGs from Nova Scotia branch with toxigenic and non-toxigenic *Microcoleus* genomes retrieved from New Brunswick (Canada), New Zealand and the USA (Bouma-Gregson et al. 2019; Conklin et al. 2020; Tee et al. 2021; Valadez-Cano, Reyes-Prieto, Beach et al. 2023), along with the recently reported toxigenic *Tychonema bourrellyi* B0820 (Salmaso et al. 2023) (Figure 2). With the exception of *Microcoleus* sp. NSFL-A1, which was recovered from a small stream connecting Fish Lake to Grand Lake, the remaining eight of these *Microcoleus* cMAGs from Nova Scotia harbour genes responsible for the production of ATXs (Figure 2). The ML phylogenetic analysis recovered the two *Microcoleus* clades as part of a larger monophyletic group (100% BS) branching with full node support as siblings of the toxigenic *Kamptonema* spp., which are also members of the Microcoleaceae family (Figure 2).

The ANI and dDDH results (Tables S5 and S6, respectively) suggest that the cMAGs recovered from Nova Scotia belong to 14 putative cyanobacterial species clusters when applying standard species delimitation cutoffs (SCs) (Table S7). Among these, the 16 *Microcoleus* cMAGs were grouped into six distinct SCs. The

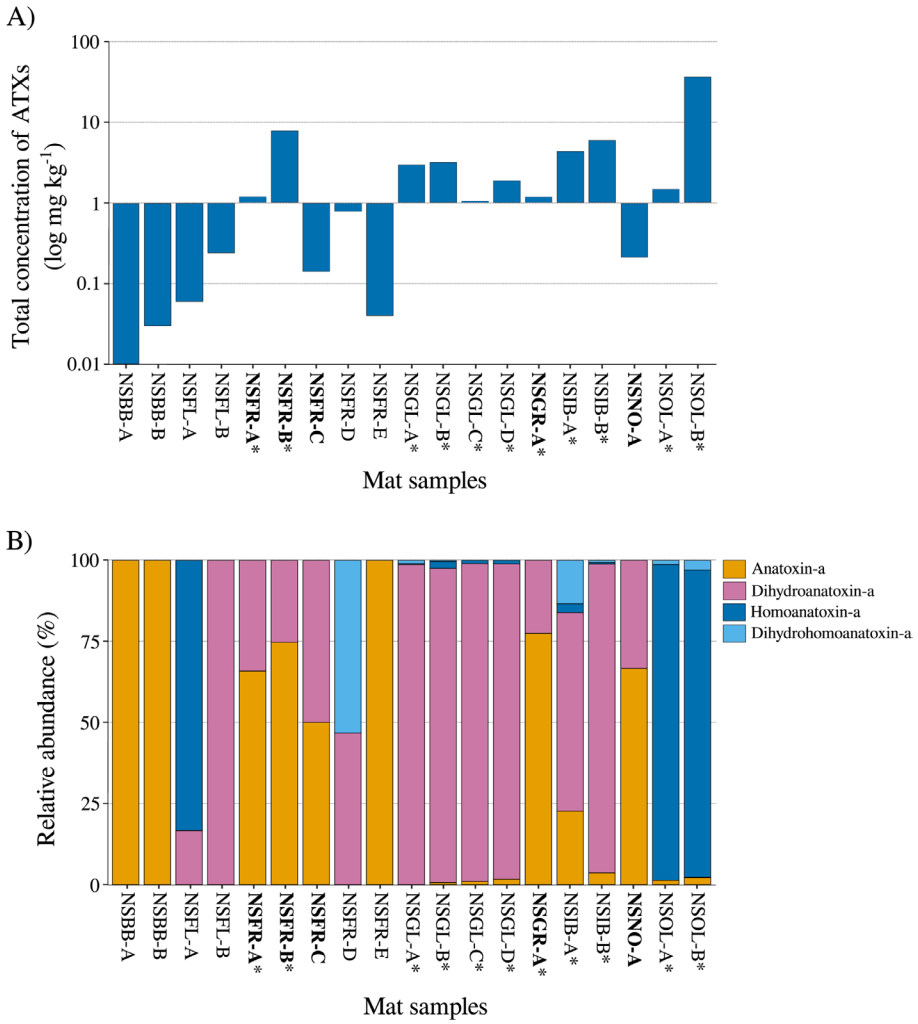

FIGURE 1 | Total concentration and relative abundance of ATXs in benthic mats collected from Nova Scotia. (A) Log_{10} of total concentrations ($mg\,kg^{-1}$) of ATXs in mat samples from Nova Scotia. (B) Relative abundance of anatoxin-a and its analogues in the different mat samples. Sample names from Nova Scotia (NS) include site codes indicating sampling locations: Byers Brook (BB), Fish Lake (FL), French River (FR), Grand Lake (GL), Great Village River (GR), Indian Brook (IB), North River (NO) and Oathill Lake (OL). Mat sample NSLS-A, with no ATXs detected, has been omitted. Samples collected in 2022 are indicated in bold. Mats with total concentration of ATXs > $1\,mg\,kg^{-1}$ are indicated with '*'.

cMAGs containing the *ana* cluster in SC8 exhibited ANI values above 95% and dDDH values exceeding 70% with both toxigenic and non-toxigenic *Microcoleus* from the Wolastoq (Valadez-Cano, Reyes-Prieto, Beach et al. 2023) and New Zealand (Tee et al. 2021) (Table S7). The remaining *Microcoleus* cMAGs recovered from Nova Scotia were assigned to five distinct SCs, with ANI and dDDH values below the thresholds for species delimitation when compared to previously reported *Microcoleus* genomes (Table S7). *Microcoleus* SC9 includes two cMAGs recovered from the French River during both sampling years containing the *ana* cluster (Table S7). Finally, the eight cMAGs classified in the Coleofasciculaceae family were assigned to seven putative novel cyanobacterial species (Table S7).

3.3 | Species Diversity and Distribution in Aquatic Systems of Nova Scotia

We employed a sequencing coverage (i.e., read mapping) approach to account for the presence of *Microcoleus* species and other cyanobacterial SCs that may exist at low abundances, and thus the assembly of their genomes would be hindered. Our analysis revealed that *Microcoleus* dominated most of the sampled mats, accounting for over 45% of the relative abundance of the cyanobacterial community (Figure 3A). Only four samples were dominated by other cyanobacterial species with relative abundances higher than 95% of the cyanobacterial community (Figure 3A).

Consistent with our earlier observations, the *Microcoleus* community is dominated by species of either MC1 or MC2 (Figure 3B). As expected, the dominant *Microcoleus* SCs are primarily composed of cMAGs from local aquatic systems in Nova Scotia (SC1, SC3, SC6, SC8, SC9, SC13). However, *Microcoleus* species members of MC2 and recovered from geographically distant aquatic systems comprised between 1% and 5% of the total *Microcoleus* community (Figure 3B). For example, we detected SC11 and SC12, which include *M. anatoxicus* from California and the non-toxigenic *Microcoleus* EPA2 recovered from Ohio, USA (Figure 3B). We also identified the presence of SC15, which includes both toxigenic and non-toxigenic *Tychonema bourrelyi* strains from Italy, although in very low abundance (Figure 3B).

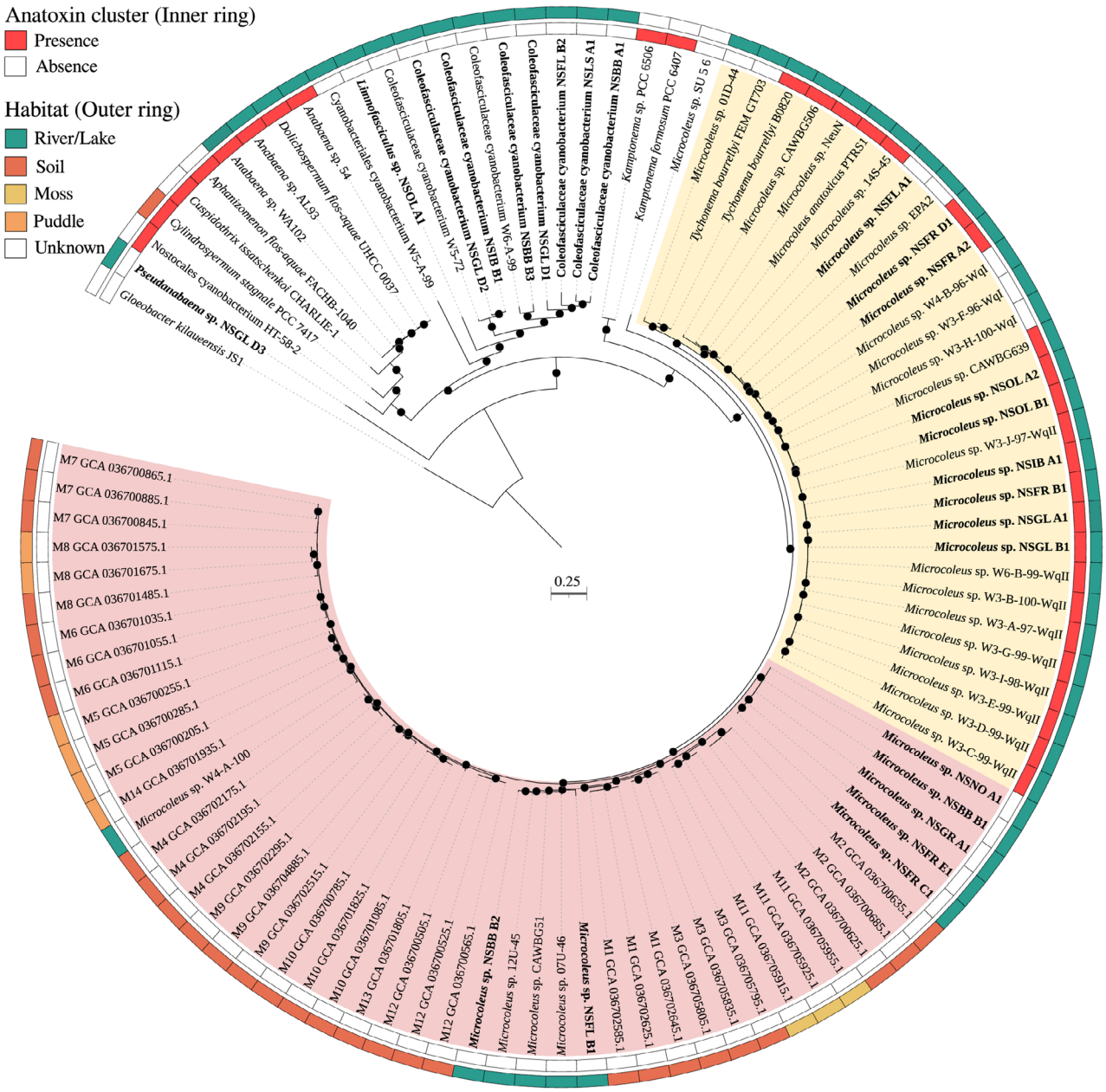

FIGURE 2 | Maximum likelihood phylogeny estimated using the Q.plant + F + I + R6 substitution model based on a concatenated alignment (data set available upon request) of 1025 orthogroups comprising 286,340 amino acid sites from *Microcoleus* and other cyanobacterial genomes. Cyanobacterial genomes recovered from Nova Scotia are indicated with bold. Clades in red and yellow indicate the two major *Microcoleus* clades 1 and 2, respectively. Inner and outer rings indicate presence of the *ana* cluster and the respective habitat from which each genome was obtained, respectively. Black circles at the nodes indicate bootstrap support values of 100%. The scale bar represents amino acid substitutions per site. The tree was rooted arbitrarily using *Gloeobacter kilaueensis* JS1 as outgroup.

In mats dominated by *Microcoleus* from MC1, we observed three different SCs, with SC1 being dominant (Figure 3B). In mats dominated by MC2, we identified eight SCs, with SC8 being the most prevalent and comprised of cMAGs with or without the *ana* cluster (Figure 3B). The relative abundance of SC8 exceeded 50% of the *Microcoleus* community in nine mat samples collected from lentic or lotic systems, with total intracellular ATXs > 1 mg kg^{-1} in benthic mats (Figure 3B). SC9, which contains the novel toxigenic *Microcoleus* species, dominated (relative abundances > 95%) two samples from French River. The non-*Microcoleus* cyanobacterial community was predominantly composed of a single species in most samples, with no clear dominance observed (Figure 3C).

The estimated Shannon index revealed a higher cyanobacterial diversity in samples collected from lentic systems compared to those from lotic systems (Wilcoxon test, $p < 0.05$, $W = 71$) (Figure S3). On the other hand, no significant differences were observed in samples collected from different ecoregions, substrates and trophic states. The Jensen–Shannon divergence analysis suggested that substrate ($R^2 = 33\%$, $p = 0.01$) and ecoregion ($R^2 = 22\%$, $p = 0.006$) were the most influential variables explaining the

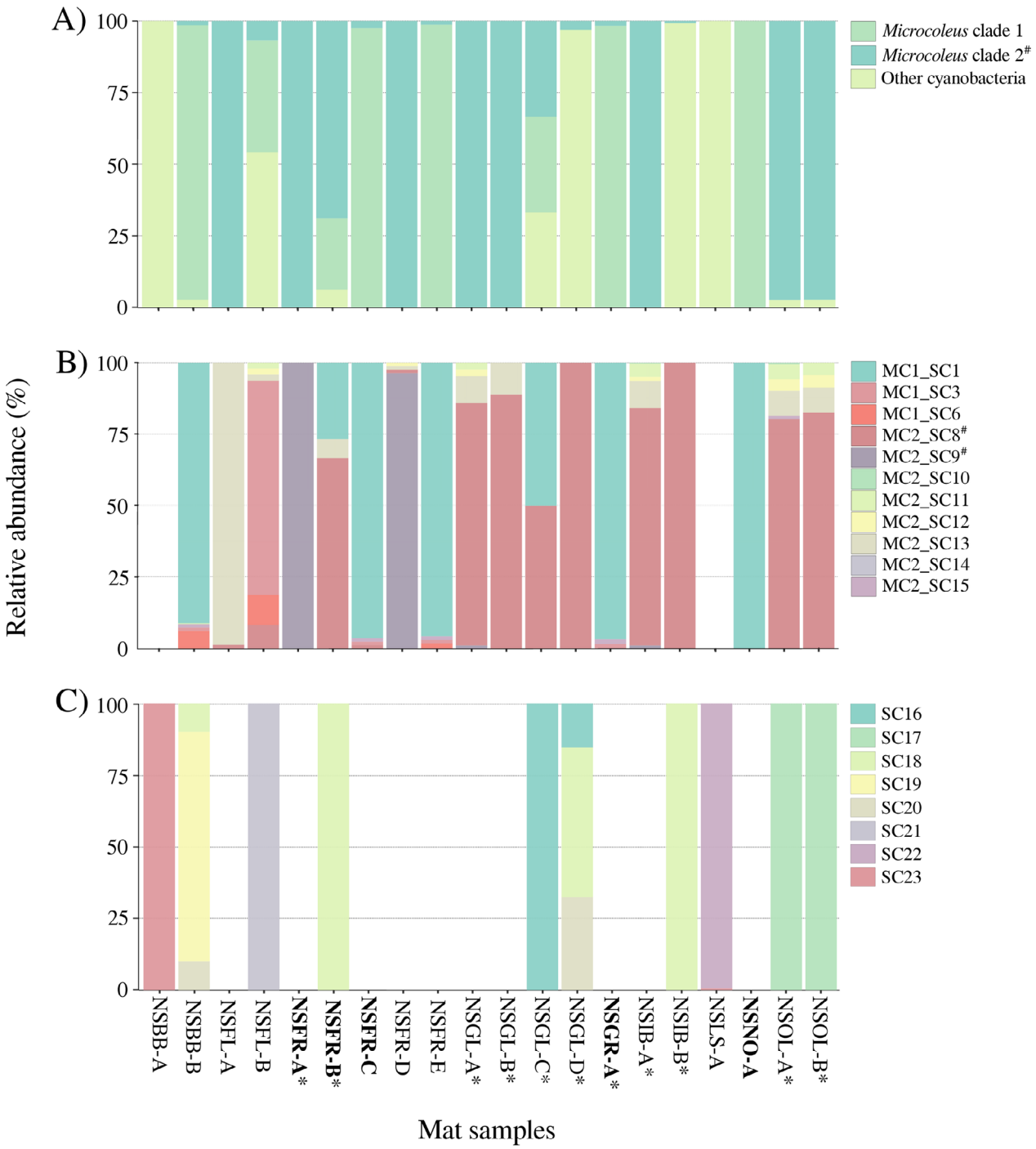

FIGURE 3 | Estimated relative abundance of *Microcoleus* and other cyanobacteria in freshwater benthic mats from Nova Scotia. (A) Relative abundance of *Microcoleus* clades 1 and 2 and other cyanobacteria. (B) Relative abundance of the different species clusters (SC) assigned to *Microcoleus* clade 1 (MC1) and *Microcoleus* clade 2 (MC2). (C) Relative abundance of non-*Microcoleus* cyanobacterial species (SC). Clades and species clusters containing *Microcoleus* cMAGs with the *ana* cluster are marked with '#' in panels (A, B), respectively. Mat samples collected in 2022 are indicated in bold, and samples with total concentrations of ATXs > 1 mg kg^{-1} are denoted with an asterisk ('*').

cyanobacterial composition in the mats. This result was confirmed by the non-parametric *ANOSIM* test (Table 1).

Our redundancy analysis did not reveal any obvious association between the concentration of ATXs and the abundance of particular toxigenic *Microcoleus* species, likely due to the similar toxin profiles of the dominant toxigenic species (SC8 and SC9). Interestingly, hATX in the mats was primarily associated with *Limnofasciculus* sp. NSOL-A1 (SC17) (Figure S4). The adjusted R^2 value was 0.678, with the first axis explaining 72.6% of the variation in cyanotoxins (ANOVA, $p < 0.05$).

3.4 | Variability of the *ana* Cluster Among *Microcoleus* Species

The synteny of the *ana* cluster (i.e., *anaBCDEFG*) is highly conserved in the different *Microcoleus* SCs, and the nucleotide

TABLE 1 | Environmental variables explaining cyanobacterial community variation.

Variable	R^2 (%)	p	Dispersion p	ANOSIM R	ANOSIM p
Ecoregion	22	0.006	$p > 0.05$	0.27	0.01
River or lake	14	0.009	$p > 0.05$	−0.013	0.48
Substrate	33	0.01	$p > 0.05$	0.20	0.06
Trophic state	15	0.09	$p > 0.05$	0.065	0.23
Years	9.3	0.08	$p > 0.05$	0.10	0.20

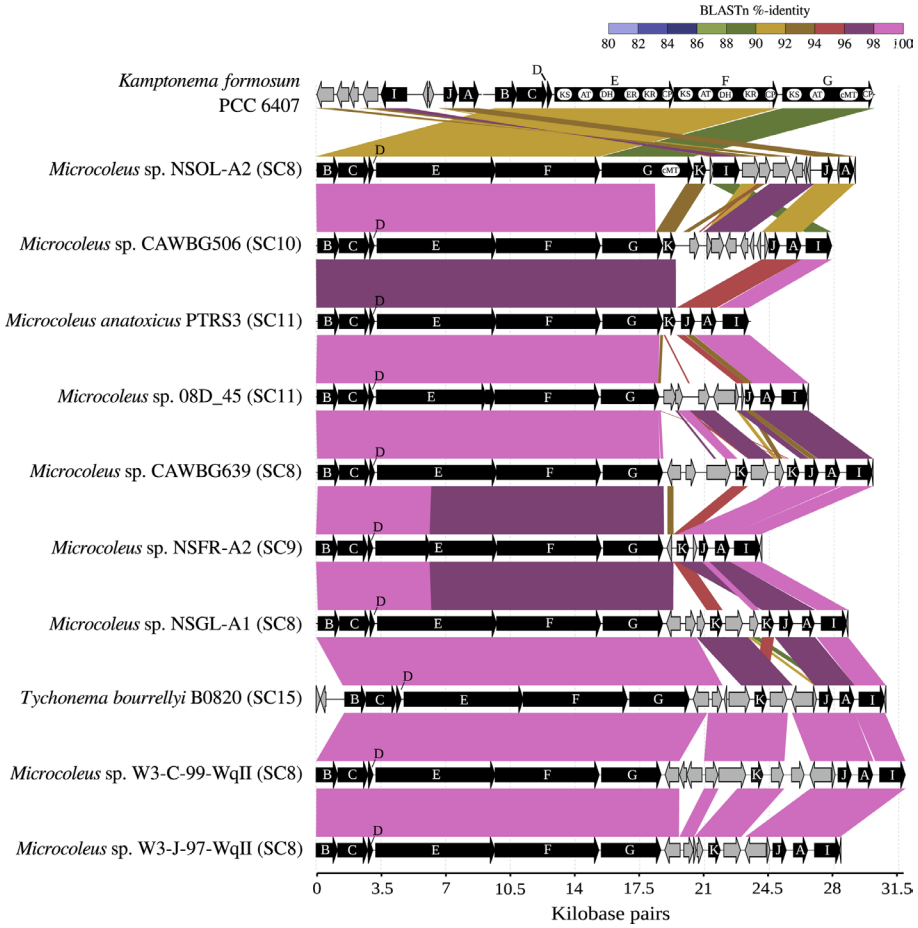

FIGURE 4 | Alignment of *ana* clusters recovered from different *Microcoleus* species clusters (SCs) and *Kamptonema formosum* PCC 6407. Black arrows indicate previously characterised genes responsible for the production of ATXs while grey arrows indicate uncharacterized ORFs. White ovals within the *anaEFG* genes from the *K. formosum* PCC 6407 *ana* cluster indicate important catalytic domains of the encoded polyketide synthases: KS: ketosynthase; AT: acyltransferase; KR: ketoreductase; DH: dehydratase; ER: enoylreductase; CP: carrier protein; cMT: methyltransferase. Diagonal lines between *ana* clusters represent the ORF nucleotide sequence identity (BLASTn percent) indicated by the scale bar. The white oval in the *anaG* gene of *Microcoleus* sp. NSOL-A2 indicates the unique insertion encoding the cMT domain, which is absent in the *ana* cluster of other members of MC2. Identical *ana* clusters were omitted.

sequence identity at the orthologous gene level is > 97% (Figure 4). Despite the high conservation of synteny between *Kamptonema* and *Microcoleus*, their orthologous nucleotide sequence identity is less than 91% (Figure 4). While there was conservation of the *anaBCDEFG* gene order, we found variability in the gene content and arrangement of the *anaAIJ* section of the *ana* cluster between different *Microcoleus* species (Figure 4).

In *ana* clusters recovered from *Microcoleus* SC8, we observed variation in the number of *anaK* copies (Figure 4). Notably, *Microcoleus* sp. NSGL-A1 contains two copies of *anaK*, which is similar to *Microcoleus* sp. CAWBG639 from New Zealand but different from other *ana* clusters in SC8 from Nova Scotia and New Brunswick, which contain a single copy (Figure 4). Additionally, the *ana* clusters in *Microcoleus* sp. NSOL-A2 and *Microcoleus* sp. NSOL-B1 from Oathill Lake differ significantly from other *Microcoleus* SC8 cMAGs, with *anaJAI* genes exhibiting less than 91% sequence identity to homologues in other SC8 *ana* clusters and *anaI* translocated relative to its position in these clusters (Figure 4). The *ana* cluster of *Microcoleus* sp.

NSOL-A2 and *Microcoleus* sp. NSOL-B1 contains a predicted *anaG* gene (4902 bp) that is longer than the orthologous sequences found in any other *ana* cluster within MC2, which typically are around 3222 bp long (Figure 4). Unlike previously reported *anaG* genes in *Microcoleus*, which encode ketoreductase (KR)-acyltransferase (AT)-carrier protein (CP) domains, the additional ~1.7 kb insertion in *Microcoleus* sp. NSOL-A2 and *Microcoleus* sp. NSOL-B1 encodes an additional methyltransferase (cMT) domain (Figure 4).

3.5 | Coexistence of *Microcoleus* Strains With Long and Short *anaG*

To explore the relative abundance of *anaG* genotypes within samples and their potential contribution to the production of ATX and hATX, we mapped high-quality reads from all mat samples to the *anaBCDEFG* genes from *Microcoleus* sp. NSOL-A2, which contains the long *anaG*. The sequence coverage data revealed that, besides the mats from Oathill Lake, sample NSIB-A (Indian Brook) also contained the methyltransferase domain in the *anaG* ORF with a nucleotide sequence identity > 98% and mean coverage of 49 reads/nucleotide (Figure 5A). *Microcoleus* sp. NSIB-A1 recovered from this mat with high relative abundance (mean coverage > 700 reads/nucleotide), harbours the short *anaG* version.

The single nucleotide variants (SNVs) identified in sample NSIB-A by read mapping to the *ana* cluster confirm the strain heterogeneity of toxigenic *Microcoleus* in this Indian Brook mat sample (Figure 5A). Furthermore, the NSIB-A sample contained a mixture of ATXs, with hATX accounting for 16% of the total analogues (Figure 1). In contrast, this heterogeneity was not observed in sample NSOL-A, where only the strain with the long *anaG* version was recovered, and the toxin profile was dominated by hATX and H_2hATX (Figure 5B).

4 | Discussion

4.1 | Geographic Prevalence of ATXs in Aquatic Systems in Nova Scotia

Following earlier reports of dog deaths linked to ATXs in Nova Scotia's aquatic systems (Crosby 2023; Johnston et al. 2024), the presence of these neurotoxins in 19 of 20 benthic mat samples over two consecutive years strongly supports the widespread environmental distribution of toxigenic microorganisms across the province. The geographical prevalence of cyanobacteria that produce ATXs has also been documented in water bodies in New Brunswick (Beach et al. 2023; McCarron et al. 2023; Valadez-Cano, Reyes-Prieto, Beach et al. 2023), suggesting that this risk likely extends to water bodies across Atlantic Canada. While several stakeholder groups have directed efforts towards monitoring cyanotoxins in planktonic environments (e.g., microcystins), our results highlight the significance of ATXs in benthic environments and the importance of monitoring the presence of these potent cyanotoxins and the cyanobacteria that produce them.

4.2 | Diversity of Mat-Forming Cyanobacteria in Nova Scotia

Our analysis revealed the presence of multiple cyanobacterial species dominating benthic mats in both lotic and lentic freshwater systems in Nova Scotia. The results showed a higher diversity of cyanobacteria in benthic mats from lentic environments compared to those from lotic environments, but this apparent difference in species richness between flowing and standing water systems must be evaluated with a larger sample size. Furthermore, our findings suggest that benthic cyanobacterial community composition is primarily influenced by geographic proximity and substrate type, highlighting

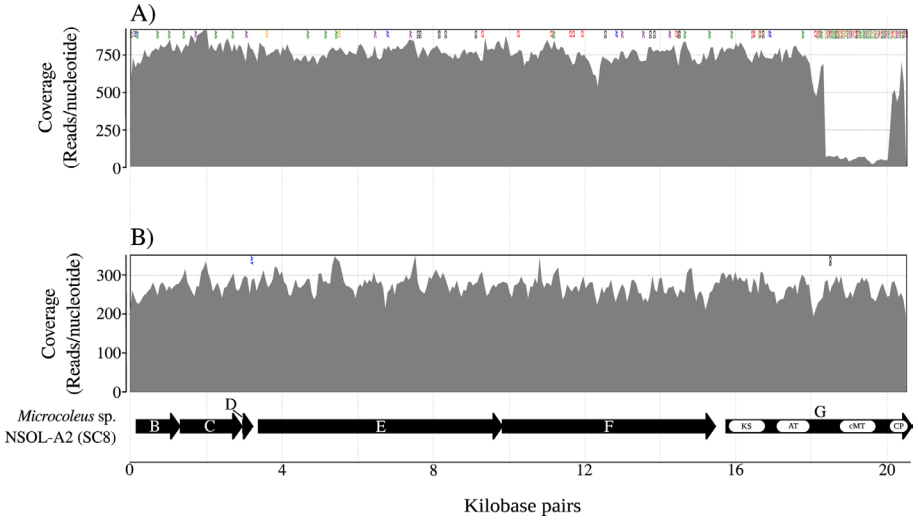

FIGURE 5 | Coexistence of toxigenic *Microcoleus* with short and long *anaG* genotypes in mat samples from Nova Scotia. Sequence coverage estimated via mapping of cleaned reads obtained from (A) Indian Brook (NSIB_A) and (B) Oathill Lake (NSOL_A) to the *anaBCDEFG* genes from *Microcoleus* sp. NSOL-A2. White ovals in the *anaG* gene from the NSOL-A2 *ana* cluster indicate the PKS catalytic modular domains important for the production of ATXs: KS: ketosynthase; AT: acyltransferase; cMT: methyltransferase; CP: Carrier protein. Dinucleotides at the top of each plot represent single nucleotide variants (SNVs).

the importance of both local (substrate characteristics) and regional (ecoregional context) factors. Substrate type likely impacts nutrient availability and colonisation surfaces, while ecoregion reflects broader climatic and biogeographical influences. These results emphasise the need to consider both spatial scales to better understand cyanobacterial diversity and toxin production in mat-forming ecosystems. With the current limited sample size, we did not detect significant differences in the cyanobacterial community composition based on the trophic states of the aquatic systems. Previous research has shown that the abundance of benthic cyanobacterial genera is influenced by a distinct set of environmental conditions, with nutrients and temperature being the less influential (Schulte et al. 2022).

The mats sampled from Nova Scotia were primarily dominated by *Microcoleus*; however, we also found other cyanobacteria, mainly from the Coleofasciculaceae family that were dominant taxa within the mats or coexisted with *Microcoleus*. Most of these Coleofasciculaceae cyanobacterial genomes from Nova Scotia clustered with genomes from isolates recovered previously from benthic mats of the Wolastoq, New Brunswick (Valadez-Cano, Reyes-Prieto, Beach et al. 2023). These results indicate that species of the Coleofasciculaceae family are a common component of mat-forming communities across these two contiguous Canadian provinces. The association of *Limnofasciculus* sp. NSOL-A1 with hATX-producing mats suggests potential ecological interactions between *Microcoleus* and other cyanobacteria within the mats. Such interactions may create favourable conditions for the synthesis of ATXs, potentially favouring strains that produce specific variants like hATX. Comparative genomic analyses have previously revealed substantial genome streamlining in toxigenic *Microcoleus* species, along with a potential reliance on external sources for nutrients, which could promote cooperative interactions between toxigenic and non-toxigenic *Microcoleus* species (Tee et al. 2021). Similarly, interactions with these co-occurring species might play an essential role in sustaining and shaping the toxigenic community within these mats.

We described the presence of at least six putative *Microcoleus* species in mats from Nova Scotia, including five novel species. This exploratory analysis suggests that *Microcoleus* SC8 is the most common species in Nova Scotia systems, as has also been described in the Wolastoq, NB (Canada) and New Zealand (Tee et al. 2021; Valadez-Cano, Reyes-Prieto, Beach et al. 2023). The dominance of SC8 in the *Microcoleus* community correlates with high concentrations of ATXs in the sampled mats, suggesting that the relative abundance of this species could serve as an indicator of mat toxicity. Additionally, we report the novel *Microcoleus* SC9, which has the genetic potential to produce ATXs. Unlike SC8, *Microcoleus* SC9 was identified in French River exclusively and during both sampling years. This sampling site lies within a river section that receives backwash discharge from a drinking water treatment plant using ultrafiltration and nanofiltration. The backwash, generated during filter cleaning, concentrates natural river constituents such as nutrients and major/minor elements, creates a unique geochemical environment. The distinctive geochemistry of this site, combined with the potential prevalence of *Microcoleus* sp. SC9, warrants further investigation to better understand their ecological roles and potential implications.

Although some mats are completely dominated by non-toxigenic cyanobacteria, we detected ATXs in most of them even when no toxin-producing species were detected. This suggests that toxigenic species can persist in the environment at very low abundances while actively producing ATXs. Alternatively, while the methods used in this work targeted intracellular toxins, low levels of ATXs may have been due to high concentrations of the toxin in the surrounding water.

4.3 | *Microcoleus* Diversity and *ana* Cluster Evolution

The phylogenomic analysis resolved *Microcoleus* species into two distinct clades, each associated with different environmental niches (Figure 2). MC1 includes species recovered from diverse habitats including soil, moss, puddles, rivers and lakes. The most prominent representative of MC1 is *Microcoleus vaginatus*, a cosmopolitan and abundant terrestrial cyanobacterium commonly found in soil crusts of arid and semiarid ecosystems (Belnap 2013; Stanojković et al. 2024). MC2 includes *Microcoleus* species predominantly recovered from lentic and lotic systems across diverse geographic regions (Bouma-Gregson et al. 2019; Tee et al. 2021; Valadez-Cano, Reyes-Prieto, Beach et al. 2023; Junier et al. 2024), suggesting that MC2 species are potentially restricted to aquatic environments.

Another notable difference between the two *Microcoleus* clades is their distinct ability to produce ATXs. Most species within *Microcoleus* clade 2, including those recovered from Nova Scotia, harbour the *ana* gene cluster. The high similarity among the *anaBCDEFG* genes within this cluster supports the hypothesis of a single acquisition event in this clade (Junier et al. 2024). Although this single acquisition event is proposed to have occurred after the divergence of *Microcoleus* from *Tychonema* (Junier et al. 2024), the *ana* cluster in *Tychonema bourrellyi* strain B0820 is nearly identical to that observed in *Microcoleus* sp. Wq-II from the Wolastoq. This contrasts with the placement of these taxa in our whole-genome phylogenetic analysis and the 87% ANI between their genomes. These findings suggest the possibility of an independent horizontal gene transfer event leading to the acquisition of the *ana* gene cluster in *Tychonema*.

The absence of the *ana* cluster in *Microcoleus* SC13 from Nova Scotia, which belongs to the MC2 clade, is consistent with previous reports of non-toxigenic species within this group. Analysis of the flanking regions of the *ana* cluster and its homologues in non-toxigenic *Microcoleus* suggests that the loss of neurotoxin production in MC2 is the result of multiple independent losses of the gene cluster, rather than repeated acquisitions (Tee et al. 2021; Junier et al. 2024). Environmental conditions may favour the loss of *ana* clusters, as secondary metabolites can impose energetic costs (Yan et al. 2018). Non-toxigenic *Microcoleus* have been found to reach higher growth rates, suggesting a trade-off between the production of ATXs and metabolic efficiency (Heath et al. 2016).

4.4 | First Description of an *anaG* Coding a Methyltransferase in *Microcoleus*

We report the first identification of an *ana* cluster in *Microcoleus* that includes a longer variant of the *anaG* gene encoding a methyltransferase domain. This extended *anaG* in *Microcoleus* is approximately 4900 bp in length and shares 88.5% pairwise nucleotide identity with the *anaG* reported in the hATX-producing *Kamptonema* sp. PCC 6506 (Méjean et al. 2009). Although it has been proposed that additional factors may regulate the activity of the *anaG* methylation domain during ATX or hATX biosynthesis in *Kamptonema* (Aráoz et al. 2005; Mann et al. 2012; Méjean et al. 2014), our results indicate that the predominant production of hATX, with minimal ATX detected in mats NSOL-A and NSOL-B, correlates with the presence of the methyltransferase domain in the extended *anaG* in *Microcoleus*. However, the regulatory mechanisms underlying the preferential production of ATX or hATX in *Microcoleus* require further investigation.

Similar to the length variation of *anaG* observed in *Microcoleus*, a comparable length variation is noted in *Cuspidothrix issatschenkoi*, where strain CHABD3 has lost the gene fragment encoding the methyltransferase domain (Jiang et al. 2015). The high nucleotide identity shared between the *anaEF* genes and the KR-AT-CP domains encoded by *anaG* across *Microcoleus ana* clusters suggests that their common ancestor likely possessed the longer *anaG* variant. This implies that the loss of the methyltransferase-encoding fragment may have occurred in most reported toxigenic *Microcoleus* genomes. Alternatively, the fragment encoding the methyltransferase could have been acquired independently in this specific *Microcoleus* strain. In either case, the evolutionary mechanisms driving the gain, loss, or retention of the methyltransferase domain in *anaG*, and thus influencing the differential production of methylated analogues, require further study.

4.5 | Anatoxin-a and Analogue Profiles From Aquatic Systems From Nova Scotia

Toxin analysis revealed a heterogeneous profile of ATXs across *Microcoleus*-dominated mats from Nova Scotia. Environmental factors such as salinity stress and variations in nitrogen and phosphorus concentrations can influence the differential production of these analogues (Heath et al. 2014; Stancheva et al. 2024). However, the variability we detected within the *ana* clusters in Nova Scotia samples suggests that spatial distribution and potential coexistence of *Microcoleus* strains in aquatic systems may also contribute to differences in ATXs ratios. While culture-based approaches are essential to confirm the production of specific ATXs by strains with distinct *ana* clusters, the comparison of gene clusters identified in environmental samples—where diverse mixtures of strains are present—with measured toxin concentrations offers valuable insights into the ecological and genetic factors influencing toxin production.

In Oathill Lake mats, which predominantly contained hATX, we identified a *Microcoleus* strain containing the long *anaG* gene version encoding the methyltransferase domain, possibly responsible for hATX production (Méjean et al. 2009). In contrast, the *Microcoleus* MAG recovered from the NSIB-A mat contained a truncated *anaG* gene lacking the methyltransferase domain, which does not account for the 16% contribution of hATX and H_2hATX to the total ATXs ($>4\,mg\,kg^{-1}$) observed in this sample. Our sequence coverage analysis revealed the presence of a strain with the long *anaG* gene at lower abundance, potentially explaining the observed ATXs mixture, although the mechanisms underlying the preferential production of H_2hATX over hATX remain unclear. In mat samples NSFL-A and NSFR-D, hATX and H_2hATX were also detected, but the strain with the long *anaG* variant was not detected. Notably, these mats contained significantly lower concentration of ATXs (0.06 and $0.78\,mg\,kg^{-1}$, respectively), suggesting that the strain with the long *anaG* variant may be present at low abundance, eluding detection by our sequencing methods. The consistent production of hATX in diverse *Microcoleus*-dominated mats from New Zealand has been documented (Heath et al. 2011; Wood et al. 2012; Kelly et al. 2018); however, genetic evidence of the *ana* cluster in these dominating *Microcoleus* remains lacking, preventing verification of the presence or absence of strains with the methyltransferase-encoding *anaG*.

Our analysis also revealed the presence of mats in Nova Scotia that predominantly produce either ATX or H_2ATX (Figure 1). Unlike *Microcoleus anatoxicus*, which has been found to predominantly produce H_2ATX in both mats and isolates (Conklin et al. 2020; Junier et al. 2024; Stancheva et al. 2024), *Microcoleus* sp. Wq-II, isolated from the Wolastoq, primarily produced ATX (Valadez-Cano, Reyes-Prieto, Beach et al. 2023). H_2ATX production has been linked to the *anaK* gene (Méjean et al. 2009, 2014, 2016), which has been identified in all known *Microcoleus ana* clusters, except a strain of *M. anatoxicus* recovered from the Eel River (Bouma-Gregson et al. 2019; Conklin et al. 2020; Tee et al. 2021; Valadez-Cano, Reyes-Prieto, Beach et al. 2023; Junier et al. 2024). This suggests that, beyond the presence or absence of *anaK*, other factors may influence the production ratio of ATX and H_2ATX. Environmental conditions likely influence ATX and H_2ATX production, as observed under nitrogen-depleted conditions in *M. anatoxicus* (Stancheva et al. 2024). However, our analysis also revealed strains in Nova Scotia containing either one or two copies of the *anaK* gene (Figure 4). The impact of possessing a single or double copy of the *anaK* gene on the differential production of ATX and H_2ATX should be considered in future studies.

Variations in the concentrations of various ATXs may arise from differential gene regulation in response to environmental conditions, but genetic heterogeneity within *Microcoleus* could also be a key factor. This includes strains differing in the presence or absence of the methyltransferase domain in *anaG*, possessing single or double copies of *anaK*, or entirely lacking *anaK*, potentially coexisting within the same mats. As shown in previous studies where variability in the concentration of ATXs between mats was largely attributed to the relative abundance of toxigenic and non-toxigenic genotypes (Wood and Puddick 2017), it is plausible that the ratios of anatoxin-a and its analogues are influenced by the differential proliferation of strains with distinct *ana* clusters. Given the phylogenetic proximity of toxigenic *Microcoleus*, which are morphologically indistinguishable and share high 16S rRNA sequence similarity, accounting for these cluster variations will be critical for future research.

5 | Conclusions

In this study, we describe the diversity and distribution of mat-forming cyanobacteria in different freshwater systems of Nova Scotia and find ATXs prevalent in most sampled lentic and lotic environments. The widespread occurrence of these cyanotoxins underscores the potential risks to both wildlife and human health in the rivers and lakes of the province. We uncovered a rich species diversity of benthic cyanobacteria in Nova Scotia freshwater mats, with members of the genus *Microcoleus* as the dominant taxa, including species with the potential to produce ATXs. This highlights both the extensive intraspecific diversity and the significant potential for the production of ATXs within this genus. Our analysis also revealed that a *Microcoleus* species present in Nova Scotia contributes substantially to the overall production of ATXs. We observed variability of the *ana* cluster, including the first report of the long *anaG* encoding a methyltransferase domain potentially responsible for hATX production in *Microcoleus*. These findings emphasise the need for continued monitoring and a deeper understanding of the ecological drivers underlying cyanotoxin production in benthic environments, which are critical for mitigating the risks associated with these harmful toxins.

Author Contributions

Cecilio Valadez-Cano: conceptualization, investigation, writing – original draft, methodology, visualization, formal analysis, data curation, writing – review and editing. **Nicolas Tromas:** methodology, writing – original draft, formal analysis, data curation, writing – review and editing. **Adrian Reyes-Prieto:** conceptualization, writing – review and editing, formal analysis, methodology. **Lindsay Johnston:** methodology, formal analysis. **Yannan Huang:** methodology, formal analysis. **Hannah Morris:** methodology, formal analysis. **Lydia Zamlynny:** methodology, formal analysis. **Daniel G. Beach:** methodology, writing – review and editing, formal analysis. **Rob C. Jamieson:** methodology, writing – review and editing, formal analysis, funding acquisition, supervision. **Janice Lawrence:** conceptualization, investigation, writing – review and editing, funding acquisition, supervision.

Acknowledgements

Funding for this study was provided by the Natural Sciences and Engineering Research Council of Canada (ALLRP 576532—22).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data generated in this study are publicly available. All sequence data have been deposited with NCBI under the Bioproject PRJNA1162496. Supplementary figures and tables have been deposited to Figshare with the following DOI: https://doi.org/10.6084/m9.figshare.28344317.v2.

References

Anderson, M. J. 2001. "A New Method for Non-Parametric Multivariate Analysis of Variance." *Austral Ecology* 26: 32–46.

Anderson, M. J. 2006. "Distance-Based Tests for Homogeneity of Multivariate Dispersions." *Biometrics* 62: 245–253.

Aráoz, R., H.-O. Nghiêm, R. Rippka, N. Palibroda, N. T. de Marsac, and M. Herdman. 2005. "Neurotoxins in Axenic Oscillatorian Cyanobacteria: Coexistence of Anatoxin-a and Homoanatoxin-a Determined by Ligand-Binding Assay and GC/MS." *Microbiology (Reading)* 151: 1263–1273.

Bankevich, A., S. Nurk, D. Antipov, et al. 2012. "SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing." *Journal of Computational Biology* 19: 455–477.

Beach, D., L. Zamlynny, M. MacArthur, and C. Miles. 2023. "Liquid Chromatography–High-Resolution Tandem Mass Spectrometry of Anatoxins, Including New Conjugates and Reduction Products." *Analytical and Bioanalytical Chemistry* 415: 5281–5296.

Beach, D. G., M. Bruce, J. Lawrence, and P. McCarron. 2022. "Rapid Quantitation of Anatoxins in Benthic Cyanobacterial Mats Using Direct Analysis in Real-Time–High-Resolution Tandem Mass Spectrometry." *Environmental Science & Technology* 56: 13837–13844.

Belnap, J. 2013. "Crusts|Biological." In *Reference Module in Earth Systems and Environmental Sciences*. Elsevier.

Blin, K., S. Shaw, A. M. Kloosterman, et al. 2021. "antiSMASH 6.0: Improving Cluster Detection and Comparison Capabilities." *Nucleic Acids Research* 49: W29–W35.

Bouma-Gregson, K., M. R. Olm, A. J. Probst, K. Anantharaman, M. E. Power, and J. F. Banfield. 2019. "Impacts of Microbial Assemblage and Environmental Conditions on the Distribution of Anatoxin-a Producing Cyanobacteria Within a River Network." *ISME Journal* 13: 1618–1634.

Cáceres, M. D., and P. Legendre. 2009. "Associations Between Species and Groups of Sites: Indices and Statistical Inference." *Ecology* 90: 3566–3574.

Chaumeil, P.-A., A. J. Mussig, P. Hugenholtz, and D. H. Parks. 2022. "GTDB-Tk v2: Memory Friendly Classification With the Genome Taxonomy Database." *Bioinformatics* 38: 5315–5316.

Chen, S., Y. Zhou, Y. Chen, and J. Gu. 2018. "Fastp: An Ultra-Fast All-in-One FASTQ Preprocessor." *Bioinformatics (Oxford, England)* 34: i884–i890.

Clarke, K. R. 1993. "Non-Parametric Multivariate Analyses of Changes in Community Structure." *Australian Journal of Ecology* 18: 117–143.

Comeau, A. M., and G. V. Filloramo. 2023. "Preparing Multiplexed WGS/MetaG Libraries With the Illumina DNA Prep Kit for the Illumina NextSeq or MiSeq." Protocols.io.

Conklin, K. Y., R. Stancheva, T. G. Otten, et al. 2020. "Molecular and Morphological Characterization of a Novel Dihydroanatoxin-a Producing *Microcoleus* Species (Cyanobacteria) From the Russian River, California, USA." *Harmful Algae* 93: 101767.

Crosby, M. 2023. "Vet Hospital Says Dog Died From Blue-Green Algae Exposure After 5 Minutes in Halifax-Area Lake." CBC News.

Edwards, C., K. A. Beattie, C. M. Scrimgeour, and G. A. Codd. 1992. "Identification of Anatoxin-a in Benthic Cyanobacteria (Blue-Green Algae) and in Associated Dog Poisonings at Loch Insh, Scotland." *Toxicon* 30: 1165–1175.

Emms, D. M., and S. Kelly. 2015. "OrthoFinder: Solving Fundamental Biases in Whole Genome Comparisons Dramatically Improves Orthogroup Inference Accuracy." *Genome Biology* 16: 157.

Emms, D. M., and S. Kelly. 2019. "OrthoFinder: Phylogenetic Orthology Inference for Comparative Genomics." *Genome Biology* 20: 238.

Eren, A. M., E. Kiefl, A. Shaiber, et al. 2021. "Community-Led, Integrated, Reproducible Multi-Omics With Anvi'o." *Nature Microbiology* 6: 3–6.

Faassen, E. J., L. Harkema, L. Begeman, and M. Lurling. 2012. "First Report of (Homo)anatoxin-a and Dog Neurotoxicosis After Ingestion of Benthic Cyanobacteria in the Netherlands." *Toxicon* 60: 378–384.

Gugger, M., S. Lenoir, C. Berger, et al. 2005. "First Report in a River in France of the Benthic Cyanobacterium *Phormidium favosum* Producing Anatoxin-a Associated With Dog Neurotoxicosis." *Toxicon* 45: 919–928.

Gunn, G. J., A. G. Rafferty, G. C. Rafferty, et al. 1992. "Fatal Canine Neurotoxicosis Attributed to Blue-Green Algae (Cyanobacteria)." *Veterinary Record* 130: 301–302.

Heath, M., S. A. Wood, R. G. Young, and K. G. Ryan. 2016. "The Role of Nitrogen and Phosphorus in Regulating *Phormidium* Sp. (Cyanobacteria) Growth and Anatoxin Production." *FEMS Microbiology Ecology* 92, no. 3: fiw021. https://doi.org/10.1093/femsec/fiw021.

Heath, M. W., S. A. Wood, R. F. Barbieri, R. G. Young, and K. G. Ryan. 2014. "Effects of Nitrogen and Phosphorus on Anatoxin-a, Homoanatoxin-a, Dihydroanatoxin-a and Dihydrohomoanatoxin-a Production by *Phormidium autumnale*." *Toxicon* 92: 179–185.

Heath, M. W., S. A. Wood, and K. G. Ryan. 2011. "Spatial and Temporal Variability in *Phormidium* Mats and Associated Anatoxin-a and Homoanatoxin-a in Two New Zealand Rivers." *Aquatic Microbial Ecology* 64: 69–79.

Hyatt, D., G.-L. Chen, P. F. Locascio, M. L. Land, F. W. Larimer, and L. J. Hauser. 2010. "Prodigal: Prokaryotic Gene Recognition and Translation Initiation Site Identification." *BMC Bioinformatics* 11: 119.

Jain, C., L. M. Rodriguez-R, A. M. Phillippy, K. T. Konstantinidis, and S. Aluru. 2018. "High Throughput ANI Analysis of 90K Prokaryotic Genomes Reveals Clear Species Boundaries." *Nature Communications* 9: 5114.

Jiang, Y., G. Song, Q. Pan, Y. Yang, and R. Li. 2015. "Identification of Genes for Anatoxin-a Biosynthesis in *Cuspidothrix issatschenkoi*." *Harmful Algae* 46: 43–48.

Johnston, L., D. Dunnington, M. Greenwood, B. Kurylyk, and R. Jamieson. 2022. "Identifying Hydrologic Regimes and Drivers in Nova Scotia, Canada: Catchment Classification Efforts for a Data-Limited Region." *Journal of Hydrologic Engineering* 27, no. 11: 05022017. https://doi.org/10.1061/(ASCE)HE.1943-5584.0002200.

Johnston, L. H., Y. Huang, T. D. Bermarija, et al. 2024. "Proliferation and Anatoxin Production of Benthic Cyanobacteria Associated With Canine Mortalities Along a Stream-Lake Continuum." *Science of the Total Environment* 917: 170476.

Junier, P., G. Cailleau, M. Fatton, et al. 2024. "A Cohesive *Microcoleus* Strain Cluster Causes Benthic Cyanotoxic Blooms in Rivers Worldwide." *Water Research X* 24: 100252. https://doi.org/10.1016/j.wroa.2024.100252.

Kalyaanamoorthy, S., B. Q. Minh, T. K. F. Wong, A. von Haeseler, and L. S. Jermiin. 2017. "ModelFinder: Fast Model Selection for Accurate Phylogenetic Estimates." *Nature Methods* 14: 587–589.

Kang, D. D., J. Froula, R. Egan, and Z. Wang. 2015. "MetaBAT, an Efficient Tool for Accurately Reconstructing Single Genomes From Complex Microbial Communities." *PeerJ* 3: e1165.

Kelly, L. T., S. A. Wood, T. G. McAllister, and K. G. Ryan. 2018. "Development and Application of a Quantitative PCR Assay to Assess Genotype Dynamics and Anatoxin Content in *Microcoleus autumnalis*-Dominated Mats." *Toxins (Basel)* 10, no. 11: 431. https://doi.org/10.3390/toxins10110431.

Langmead, B., and S. Salzberg. 2012. "Fast Gapped-Read Alignment With Bowtie 2." *Nature Methods* 9: 357–359.

Mann, S., M. Cohen, F. Chapuis-Hugon, et al. 2012. "Synthesis, Configuration Assignment, and Simultaneous Quantification by Liquid Chromatography Coupled to Tandem Mass Spectrometry, of Dihydroanatoxin-a and Dihydrohomoanatoxin-a Together With the Parent Toxins, in Axenic Cyanobacterial Strains and in Environmental Samples." *Toxicon* 60: 1404–1414.

McCarron, P., C. Rafuse, S. Scott, et al. 2023. "Anatoxins From Benthic Cyanobacteria Responsible for Dog Mortalities in New Brunswick, Canada." *Toxicon* 227: 107086.

McMurdie, P. J., and S. Holmes. 2013. "Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data." *PLoS One* 8: 1–11.

Meier-Kolthoff, J. P., A. F. Auch, H.-P. Klenk, and M. Göker. 2013. "Genome Sequence-Based Species Delimitation With Confidence Intervals and Improved Distance Functions." *BMC Bioinformatics* 14: 60.

Méjean, A., K. Dalle, G. Paci, et al. 2016. "Dihydroanatoxin-a Is Biosynthesized From Proline in *Cylindrospermum stagnale* PCC 7417: Isotopic Incorporation Experiments and Mass Spectrometry Analysis." *Journal of Natural Products* 79: 1775–1782.

Méjean, A., S. Mann, T. Maldiney, G. Vassiliadis, O. Lequin, and O. Ploux. 2009. "Evidence That Biosynthesis of the Neurotoxic Alkaloids Anatoxin-a and Homoanatoxin-a in the Cyanobacterium *Oscillatoria* PCC 6506 Occurs on a Modular Polyketide Synthase Initiated by l-Proline." *Journal of the American Chemical Society* 131: 7512–7513.

Méjean, A., G. Paci, V. Gautier, and O. Ploux. 2014. "Biosynthesis of Anatoxin-a and Analogues (Anatoxins) in Cyanobacteria." *Toxicon* 91: 15–22.

Nayfach, S., B. Rodriguez-Mueller, N. Garud, and K. Pollard. 2016. "An Integrated Metagenomics Pipeline for Strain Profiling Reveals Novel Patterns of Bacterial Transmission and Biogeography." *Genome Research* 26: 1612–1625.

Neily, P., S. Basquill, E. Quigley, and K. Keys. 2017. "Ecological Land Classification for Nova Scotia. Nova Scotia Department of Natural Resources, Renewable Resources Branch." DNR Report FOR 2017-13.

Nguyen, L.-T., H. A. Schmidt, A. von Haeseler, and B. Q. Minh. 2015. "IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies." *Molecular Biology and Evolution* 32: 268–274.

Nishimura, Y., K. Yamada, Y. Okazaki, and H. Ogata. 2024. "DiGAlign: Versatile and Interactive Visualization of Sequence Alignment for Comparative Genomics." *Microbes and Environments* 39, no. 1: ME23061. https://doi.org/10.1264/jsme2.ME23061.

Nissen, J. N., J. Johansen, R. L. Allesøe, et al. 2021. "Improved Metagenome Binning and Assembly Using Deep Variational Autoencoders." *Nature Biotechnology* 39: 555–560.

Oksanen, J., F. G. Blanchet, M. Friendly, et al. 2022. "Vegan: Community Ecology Package. R Package Version 2.5-7." 1–3.

Parks, D. H., M. Chuvochina, C. Rinke, A. J. Mussig, P.-A. Chaumeil, and P. Hugenholtz. 2022. "GTDB: An Ongoing Census of Bacterial and Archaeal Diversity Through a Phylogenetically Consistent, Rank Normalized and Complete Genome-Based Taxonomy." *Nucleic Acids Research* 50: D785–D794.

Parks, D. H., M. Imelfort, C. T. Skennerton, P. Hugenholtz, and G. W. Tyson. 2015. "CheckM: Assessing the Quality of Microbial Genomes Recovered From Isolates, Single Cells, and Metagenomes." *Genome Research* 25: 1043–1055.

Quiblier, C., S. Wood, I. Echenique, M. Heath, V. Aurelie, and J.-F. Humbert. 2013. "A Review of Current Knowledge on Toxic Benthic Freshwater Cyanobacteria—Ecology, Toxin Production and Risk Management." *Water Research* 47: 5464–5479.

Rao, C. R. 1964. "The Use and Interpretation of Principal Component Analysis in Applied Research." *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002)* 26: 329–358.

Rao, C. R. 1995. "A Review of Canonical Coordinates and an Alternative to Correspondence Analysis Using Hellinger Distance." *Qüestiió (Quaderns d'Estadística i Investigació Operativa)* 19: 23–63.

Richter, M., and R. Rosselló-Móra. 2009. "Shifting the Genomic Gold Standard for the Prokaryotic Species Definition." *Proceedings of the National Academy of Sciences* 106: 19126–19131.

Rognes, T., T. Flouri, B. Nichols, C. Quince, and F. Mahé. 2016. "VSEARCH: A Versatile Open Source Tool for Metagenomics." *PeerJ* 4: e2584.

Salmaso, N., A. Boscaini, L. Cerasino, et al. 2023. "Draft Genome Sequence of the Anatoxin-a Producing Cyanobacterium *Tychonema*

bourrellyi B0820 Isolated From the Epilimnion of the Deep Alpine Lake Garda." *Microbiology Resource Announcements* 12: e00844-23.

Schulte, N. O., D. M. Carlisle, and S. A. Spaulding. 2022. "Natural and Anthropogenic Influences on Benthic Cyanobacteria in Streams of the Northeastern United States." *Science of the Total Environment* 826: 154241.

Sieber, C. M. K., A. J. Probst, A. Sharrar, et al. 2018. "Recovery of Genomes From Metagenomes via a Dereplication, Aggregation and Scoring Strategy." *Nature Microbiology* 3: 836–843.

Stancheva, R., S. Brown, G. L. Boyer, et al. 2024. "Effect of Salinity Stress and Nitrogen Depletion on Growth, Morphology and Toxin Production of Freshwater Cyanobacterium *Microcoleus anatoxicus* Stancheva & Conklin." *Hydrobiologia* 852: 561–574.

Stanojković, A., S. Skoupý, H. Johannesson, and P. Dvořák. 2024. "The Global Speciation Continuum of the Cyanobacterium *Microcoleus*." *Nature Communications* 15: 2122.

Tee, H. S., S. A. Wood, K. Bouma-Gregson, G. Lear, and K. M. Handley. 2021. "Genome Streamlining, Plasticity, and Metabolic Versatility Distinguish Co-Occurring Toxic and Nontoxic Cyanobacterial Strains of *Microcoleus*." *MBio* 12: e0223521.

Valadez-Cano, C., A. Reyes-Prieto, D. G. Beach, C. Rafuse, P. McCarron, and J. Lawrence. 2023. "Genomic Characterization of Coexisting Anatoxin-Producing and Non-Toxigenic *Microcoleus* Subspecies in Benthic Mats From the Wolastoq, New Brunswick, Canada." *Harmful Algae* 124: 102405. https://doi.org/10.1016/j.hal.2023.102405.

Valadez-Cano, C., A. Reyes-Prieto, and J. Lawrence. 2023. "Novel Virulent and Temperate Cyanophages Predicted to Infect *Microcoleus* Associated With Anatoxin-Producing Benthic Mats." *Environmental Microbiology* 25: 3319–3332.

Wood, S. A., and J. Puddick. 2017. "The Abundance of Toxic Genotypes Is a Key Contributor to Anatoxin Variability in *Phormidium*-Dominated Benthic Mats." *Marine Drugs* 15, no. 10: 307. https://doi.org/10.3390/md15100307.

Wood, S. A., J. Puddick, R. Fleming, and A. H. Heussner. 2017. "Detection of Anatoxin-Producing *Phormidium* in a New Zealand Farm Pond and an Associated Dog Death." *New Zealand Journal of Botany* 55: 36–46.

Wood, S. A., F. M. J. Smith, M. W. Heath, et al. 2012. "Within-Mat Variability in Anatoxin-a and Homoanatoxin-a Production Among Benthic *Phormidium* (Cyanobacteria) Strains." *Toxins* 4: 900–912.

World Health Organization. 2020. "Cyanobacterial Toxins: Anatoxin-a and Analogues. Background Document for Development of WHO Guidelines for Drinking-Water Quality and Guidelines for Safe Recreational Water Environments, Geneva: World Health Organization; 2020 (WHO/HEP/ECH/WSH/2020.1)."

Wu, Y.-W., B. A. Simmons, and S. W. Singer. 2016. "MaxBin 2.0: An Automated Binning Algorithm to Recover Genomes From Multiple Metagenomic Datasets." *Bioinformatics* 32: 605–607.

Yan, Q., L. Dantas Lopes, B. Shaffer, et al. 2018. "Secondary Metabolism and Interspecific Competition Affect Accumulation of Spontaneous Mutants in the GacS-GacA Regulatory System in *Pseudomonas Protegens*." *MBio* 9: e01845-17.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.