Spring 4-15-2022

Identifying Molecular Markers for Early Detection of Toxic Cyanobacteria and Dinoflagellate

Shafqat F. Ehsan
St. Mary's University, sehsan@mail.stmarytx.edu

IDENTIFYING MOLECULAR MARKERS FOR EARLY DETECTION OF TOXIC CYANOBACTERIA AND DINOFLAGELLATE.

Shafqat F. Ehsan

HONORS THESIS

Presented in Partial Fulfillment of the Requirements for
Graduation from the Honors Program of
St. Mary's University
San Antonio, Texas

Approved by:

Erika Schwarz Taylor, PhD
Assistant Professor of Biological Sciences | Bioinformatics Specialist

Camille A. Langston

Camille Langston, PhD
Director, Honors Program

February 23, 2022

This thesis has

been formatted in accordance to the

author guidelines for the Journal of Phycology.

TABLE OF CONTENTS

ABSTRACT

Harmful algal Blooms (HABs) develop when algal colonies grow out of control, causing toxicity or injury to humans, fish, shellfish, marine mammals, and birds. Most HABs of public health concern in saltwater generally are caused by eukaryotic dinoflagellates and diatoms. Prokaryotic cyanobacteria are usually responsible for freshwater blooms although they can contribute to salt water and brackish blooms too. A common monitoring target of both groups is the saxitoxin-encoding genes. Saxitoxin(STX) is responsible for Paralytic shellfish poisoning, a foodborne illness developed from consumption of STX contaminated shellfish. Each cyanobacterial *SXT* gene cluster contains a set of core genes, common to all *SXT* clusters and a set of genes that vary between different clusters. In cyanobacteria, SxtA seems to be the only gene that initiates STX production and has been focused on this study. Currently most widely used monitoring methods require samples to be collected using specific or generalist sampling devices, then sorted and taxonomically identified individually, usually under the microscope, to derive biodiversity information from conventional morphological analysis. Molecular approaches are becoming increasingly popular as tools for measuring biodiversity and environmental management are improving. Targeted detection technologies, primarily based on PCR but increasingly incorporating novel probe-based methodologies, have ushered in a new era in rare species monitoring, such as ballast water surveillance. We hypothesized that there would be enough conserved yet divergent areas in the Saxitoxin gene cluster that can be accessed to create species specific probes that can detect only certain toxic species of concern in HABs. Through multiple sequence alignment, primer designing tools, and other bioinformatic analysis we focused on sxtA gene and identified a potential candidate region. This region now needs to be further studied in wet lab setting and real-world system and has the potential to be developed into a species-specific probe for early detection of HABs.

ACKNOWLEDGEMENTS

The author would like to acknowledge the Spring Research Fellowship program and Biaggini Research Scholars Program of Department of Biology, St. Mary's University for the opportunity to conduct this research and for funding the author for the work. Additionally, the author extends gratitude to his faculty mentor, Dr. Erika Schwarz Taylor, for her unconditional support and guidance throughout the project and writing of this document.

DEDICATION

This work is dedicated to my parents and faculties, who supported and guided me throughout my journey of exploration in science.

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

Harmful Algal Blooms

Algae are simple plants that range in size from microscopic single-celled organisms to large seaweeds. They generally form the base of food webs, however, on the other hand, their functions might be more menacing at times. Algae may grow out of control in the appropriate environmental conditions, and some of these "blooms" generate toxins that can kill fish, mammals, and birds, as well as cause human disease or death in severe circumstances (NOAA 2016). Collectively, these are referred to as harmful algae. Other algae are nontoxic and considered nuisance algae as they can block fish and invertebrates' gills, smother corals and submerged aquatic plants, pollute drinking water, discolor water, and deposit large, stinky heaps on beaches. Additionally, as the algae die, the bacteria that decompose them use the oxygen in the water creating hypoxic or anoxic zones, (NOAA 2016). HABs develop when algal colonies grow out of control, causing toxicity or injury to humans, fish, shellfish, marine mammals, and birds. While there are several elements that may contribute to HABs, it is unclear how these factors interact to produce an algal 'bloom'. Many algae species thrive when wind and water currents are suitable (Pitcher et al. 2010). HABs may also be connected to 'overfeeding' in some situations. This happens when nutrients (mostly phosphorus, nitrogen, and carbon) from lawns and farmlands run downriver to the sea and build up at a pace that 'overfeeds' the algae that naturally exist in the ecosystem, leading to eutrophication (NOAA 2014). Furthermore, natural phenomena such as slow water circulation, exceptionally high water temperatures, and extreme weather events such as storms, floods, and drought have also been linked to HABs (NOAA 2014).

Impact of HABs

HABs have significant impacts on our life. People frequently become ill after eating shellfish that contain toxins generated by these algae (NOAA 2014). Marine algal toxins cause more than 60,000 poisoning incidents globally each year with an associated mortality rate of 1.5% (Gill et al. 2003). HAB toxins in the air can induce breathing issues and, in some circumstances, asthma episodes in people who are sensitive (Camcho et al 2007). HABs can also be expensive in terms of financial losses. HABs currently cost the seafood, restaurant, and tourism businesses roughly $82 million per year in economic damages (NOAA 2014). It can diminish tourist industry as they can shut beaches and shellfish beds and reduce both recreational and commercial fishing catch.

In general, these microorganisms may quickly develop and create a high amount of biotoxins in the right conditions. As a result, poisoning can happen when humans swim, bathe, or farm economically sensitive species in contaminated water. Furthermore, the prevalence, frequency, and persistence of harmful algal blooms (HAB) have increased in many parts of the world over the past decades because of global climate change, eutrophication, urbanization, and modern agriculture (Cohen et al. 2009).

Organisms contributing to HABs

Most HABs of public health concern in saltwater generally are caused by eukaryotic dinoflagellates and diatoms (Fichez et al. 1992). Prokaryotic cyanobacteria are usually responsible for freshwater blooms although they can contribute to salt water and brackish blooms too (Paerl et al 2001).

Dinoflagellates, a huge and diverse group of eukaryotic algae in the marine ecosystem, are the most common source of toxins that harm humans (Gill et al. 2003). They are a monophyletic

group of primarily unicellular organisms distinguished from other groups by a series of distinct characteristics such as flagellar insertion, pigmentation, organelles, and nucleus features (Soyer-Gobillard et al. 1996). There is a broad range of cell size: small cells may be 10 μm wide by 12 μm long, and large ones may reach 400 μm in length (Soyer-Gobillard et al. 1996). Dinoflagellates exhibit two dimorphic flagella during at least one life cycle stage and have a nucleus with permanently condensed chromosomes and unique bases, which lacks histones and nucleosomes (Soyer-Gobillard et al. 1996). They can be found in freshwater, estuarine, or marine habitats (Wehr and Sheath 2003). Freshwater dinoflagellates are currently estimated to number between 250 and 300 species globally, with roughly 150 species documented from North America (Wehr and Sheath 2003). According to the Intergovernmental Oceanographic Commission (IOC), almost seventy dinoflagellate species produce toxins with cytolytic, hemolytic, hepatotoxic, or neurotoxic properties, depending on their chemical structure and conversion stage.

Until 1960, blue-green algae were the label given to the organisms which are now known as cyanobacteria (Haselkorn 2009). They were categorized as photosynthetic microbes alongside green algae, red algae, and brown algae. However, new biochemical evidence in the 1960s revealed that blue-green algae, unlike other algae, are more similar to bacteria: their peptidoglycan cell walls make them susceptible to penicillin; they have antibiotic sensitive bacterial-sized ribosomes; and they lack organelles like chloroplasts and mitochondria (Haselkorn 2009). They may be found in a wide range of biological niches, from desert to hot springs to ice-cold water (Haselkorn 2009).

Cyanobacteria are important for both their helpful and harmful properties. They are significant primary producers as well as a large source of secondary products, such as cyanotoxins, which are a group of toxic compounds (Rastogi et al. 2015). Increased anthropogenic eutrophication and global climate change have resulted in abundant proliferation of cyanobacteria

in freshwater, estuarine, and coastal ecosystems, raising severe concerns about dangerous bloom formation and surface water pollution across the world (Rastogi et al. 2015). Blooms of cyanobacteria and the buildup of numerous cyanotoxins in water bodies have serious ecological implications, posing a considerable risk to aquatic creatures and worldwide public health (Rastogi et al. 2015). Cyano toxins are thought to have a function in chemical defense mechanisms that provide cyanobacteria a competitive edge over other microorganisms or resist predation by higher trophic levels. (Berry et al. 2008)

Monitoring of HABS is crucial

Early detection of both cyanoHABs and dinoHABs through monitoring is crucial to prevent ecological and public health risks. A common monitoring target of both groups is the saxitoxin-encoding genes (Hurley et al. 2014). Despite the differences in the species, surprisingly the toxins appear to be synthesized in a similar way; stereochemistry and precursor incorporation patterns are identical for both kinds of organisms (Shimizu Y 1996).

Saxitoxin (STX) is responsible for Paralytic shellfish poisoning (PSP), a foodborne illness developed from consumption of STX contaminated shellfish (Hurley et al. 2014). Within hours of eating shellfish contaminated with toxic levels of STX, victims develop gastrointestinal distress and neurological symptoms, ranging from benign circumoral paresthesia and tingling of the extremities to ataxia, dysphagia, and changes in mental status (Gessner and Middaugh 1995). STX inhibits the development of neuronal impulses in animals by selectively blocking voltage-gated Na^+ channels in excitable cells (Murray et al. 2011). While most patients recover without intervention, respiratory paralysis and asphyxiation can occur quickly (Etheridge 2010). A feeling of "floating" or dissociation is reported by many victims (Gessner et al. 1997). There have been occurrences of hypertension and tachycardia as well (Gessner et al. 1997). Saxitoxin has no

antidotes at the moment, and medical treatment is only supportive (Etheridge 2010). Hence, the monitoring for harmful algal bloom through early detection methods are crucial to prevent ecological and public health risks.

Characteristics of Saxitoxin genes

Each cyanobacterial *SXT* gene cluster contains a set of core genes, common to all *SXT* clusters and a set of genes that vary between different clusters (Wiese et al. 2010). In cyanobacteria, SxtA seems to be the only gene that initiates STX production and has been focused on this study. The gene cluster includes four catalytic domains each of whose activities are predicted to be for a SAM-dependent methyltransferase (sxtA1), one for a GCN5-related N-acetyltransferase (sxtA2), one for an acyl carrier protein (sxtA3), and one for a class II aminotransferase (sxtA4) (Kellmann et al. 2008)

STX synthesis appears to be paraphyletic in cyanobacteria, as both STX producing (STX^+) and STX non-producing (STX^-) strains of the same species have been observed (Murray et al. 2011). This has led to the hypothesis that toxicity is an inherited feature that is passed down to only certain descendants subsequently and lost for some (Murray et al. 2011). The STX-pathways of the five cyanobacterial genera share a complement of 14 genes (sxtA–sxtI, sxtP–sxtS, and sxtU), commonly referred as the "core" genes, have been identified (Murray et al. 2011). Eight of these genes (sxtA, sxtB, sxtD, sxtG, sxtH/T, sxtI, sxtS, and sxtU) appear to be involved in the production of STX (Orr et al. 2013). Most of them appear to have come from other bacteria, specifically Proteobacteria, via horizontal gene transfers (HGTs) in cyanobacteria (Murray et al. 2011).

Saxitoxin structure and chemistry

STX has the basic structure of a trialkyl tetrahydropurine, with the NH_2 groups in positions 2 and 8 of the purine ring forming the two permanent guanidinium moieties (Schantz et al. 1975). It has two pKas of 8.22 and 11.28, which correspond to the 7,8,9 and 1,2,3 guanidinium groups (Cusick and Sayler 2013). The 1,2,3-guanidino group has a positive charge at physiological pH, but the 7,8,9-guanidino group is slightly deprotonated (Cusick and Sayler 2013).

Genera of organisms that have genes that code for saxitoxins

SXT producing dinoflagellates belong to the genera *Alexandrium*, *Gymnodinium* and *Pyrodinium* (Stüken et al. 2011) whilst production has been identified in several cyanobacterial genera including *Anabaena*, *Cylindrospermopsis*, *Aphanizomenon Planktothrix* and *Lyngbya (Christensen and Khan 2020).*

Environmental monitoring with eDNA and eRNA can aid in early detection of HABs

Currently samples must be collected using specific or generalist sampling devices (e.g., nets, electrofishing, filtering large volumes of water, sediment cores, SCUBA diving), then sorted and taxonomically identified individually, usually under the microscope, to derive biodiversity information from conventional morphological analysis (Zaiko et al. 2018). As a result, the number of samples and replicates that may be collected and evaluated is limited. Since resources are generally limited, surveillance of wide geographical areas is likely to be limited to rapid assessment surveys (Zaiko et al. 2018). Furthermore, as a result of convergent evolution, distinct species living in similar settings may become morphologically identical, while certain marine

organisms have varied forms at different developmental stages (Cohen et al. 2009). Because of these features, morphological approaches alone are ineffective in distinguishing between toxic and non-toxic species (Cohen et al. 2009).

Molecular approaches are becoming increasingly popular as tools for measuring biodiversity and improving environmental management (Liu et al. 2020). Rapid technological breakthroughs have led to the employment of a variety of molecular approaches in biosecurity applications in recent years (Zaiko et al. 2018). Environmental DNA/RNA from the environment can be extracted from a variety of sources, including soil, water, and excrement (Zaiko et al. 2018). These samples contain bulk nucleic acids, as well as dead cells and extracellular DNA/RNA, coming from the sample's living organisms (Zaiko et al. 2018).

Environmental DNA (eDNA) is DNA obtained from environmental sources without the use of a direct sampling method to sample target species (Ogram et al. 1987). DNA is expelled and accumulates in organisms' habitats when they interact with their environments, like ballast water. Targeted detection technologies, primarily based on PCR but increasingly incorporating novel probe-based methodologies, have ushered in a new era in rare species monitoring, such as ballast water surveillance (Darling and Frederick 2018).

Environmental RNA (eRNA) is emerging as a powerful alternative when distinguishing the living portion of a community is essential (Cristescu 2019). The presence of eDNA cannot fully confirm whether the organism is alive or dead (Cristescu 2019). However, presence of eRNA means that an organism is actively transcribing, informing us that the organism is most likely still alive (Cristescu 2019). Because RNA is far less stable in vitro than DNA, the use of eRNA for species identification has yet to be investigated (Cristescu 2019). Recent research suggests,

however, that RNA is widely excreted by organisms and persists long enough in the environment to reconstruct community composition and gene expression (Cristescu 2019).

Examples of where/how eDNA/eRNA could also be used

Protocols based on eDNA may enable for the collection of data on species distribution and relative abundance in a quick, cost-effective, and consistent manner for Improved Native Species Detection (Pilliod et al. 2008). eDNA offers an appealing option for aquatic inventory and monitoring projects for small, rare, secretive, and other species that are difficult to detect (Pilliod et al. 2008). When compared to electrofishing, snorkeling, and other existing field methods, there is growing evidence of enhanced species detection and catch-per-unit effort (Pilliod et al. 2008). As a result, employing eDNA to detect species could improve biodiversity assessments and give information on the status, distribution, and habitat requirements of lesser-known species (Pilliod et al. 2008).

eDNA might potentially be a useful technique for detecting aquatic invasive species early on (Pilliod et al. 2008). The use of eDNA techniques for invasive species monitoring might entail collecting water samples on a regular basis and screening them for several invasive species at once (Pilliod et al. 2008). Ballast water from boats, which is a source of introduction for many invasive species might also be tested. When a few surviving individuals recolonize the habitat, some rigorous eradication operations for invasive species fail (Pilliod et al. 2008). Methods based on eDNA and eRNA might be used to certify the elimination of all intruders (Pilliod et al. 2008).

OBJECTIVE:

We hypothesized that there would be enough conserved yet divergent areas in the Saxitoxin gene cluster that can be accessed to create species specific probes that can detect only certain toxic species of concern in harmful algal blooms.

METHODS

The National Center for Biotechnology Information's Nucleotide database was used to download publicly available SxtA gene sequences for dinoflagellate and cyanobacteria species. All sxtA sequences were mapped to the longest sxtA sequence in Geneious Prime 2022.0.2 (Kearse et al. 2012) to determine homologous regions for more targeted alignments. The mapping yielded two groups: 5′ and 3′ mapped groups and separate alignments were built for each cluster. The little overlap between the sequences made it impossible to reliably align all the sequences at the same time. Gene alignment was conducted with Geneious Prime's translation align tool and the default MUSCLE 3.8.425 (Edgar, 2004) settings. Where necessary, alignments were manually corrected by hand. The Primer3 2.3.7 (Rozen and Skaletsky 2000) in Geneious Prime, was used to identify candidate regions for molecular markers with default setting plus DNA probe parameter activated. The 5′ anchoring end of the suggested primers were then evaluated manually to check for the one that aligns with the greatest number of original clustered sequences. The top two were then aligned against the 3′ grouped sequences to check for similarity and differences and compared. This alignment was conducted with Geneious' translation align tool and the default Clustal Omega 1.2.2 (Sievers et al. 2011) settings.

RESULTS

1. Sequence information

   Sequences obtained from NCBI GenBank ranged in length from 300 bp to 3000 bp. Ninety-nine sequences mapped to the 5’ end of the sxtA full-length gene, and two hundred five sequences mapped to the 3′ end of the sxtA full-length gene. Sequences were clustered into a “5′” group and a “3′” group for further alignment.

2. Gene alignment

   a. For the 5′ group, the aligned consensus sequence length from the 99 initial sequences was 638bp with pairwise identity 89.8%. The sequences in this clustered ranged from 314bp-623bp.

   b. For the 3′ group, the aligned consensus sequence length from the 205 initial sequences was 873bp with pairwise identity 92.1%. The sequences in this clustered ranged from 257bp-780bp.

Figure 1. 3′ gene cluster alignment of partial sxtA sequences mapped against complete sequence sxtA (in no order)

3. Molecular marker candidates

The Primer design tool suggested five primers of which the following two were selected

5′ anchoring end sequence (Forward Primer)	3′ anchoring end sequence (Reverse Primer)	Binding site	Name
TGGACTACGCG GAGAACAAC	CCGGGGTCCAGTAG ATGTTG	476-497 and 813-832	476F-832R
GGACTACGCGG AGAACAACA	CTTGGCCAAGTTCA GAACGC	477-498 and 591-612	477F-612R

476F primer had percentage identity of >90% for 197 of the 204 sequences compared with and was completely identical (percentage identity of 100%) to 134 of the 204 sequences.

477F primer had percentage identity of >90% for 195 of the 204 sequences compared with and was completely identical (percentage identity of 100%) to 134 of the 204 sequences.

DISCUSSION

Many of the publicly available genomics sequences related to saxitoxin are partial sequences of uneven quality. As mentioned before, we are aware of a full set of core cluster of SXT genes but only very few have large enough samples of sequences to run a genomic analysis. Even the ones that have- sxtA being one of the most prominent ones- are inconsistant and are often incomplete sequences. This could be partially due to the certain regions within the gene being of a higher interest and partially due to high variability. Even the cloned sequences when aligned showed a 1-2bp variabilities at multiple locations. We aligned the partial sxtA sequences against complete sxtA sequence to determine true homologous regions. This led to two different clusters getting formed. The 3′ cluster was used for further analyses over the 5′ cluster since the first alignment had a pairwise identity of 92.1% which was better than the 89.8% of the latter.

Three of the five primer pairs suggested by the primer designing tool did not align well against many sequences. From evaluating the remaining two we found that both the forward primers 476 F and 477F had very similar alignment picking up most of the species almost perfectly. However, the suggested reverse primer pair for 476F, 832R, was outside the range of some of the sequences, hence, 477F might be a more suitable candidate.

There is no one-size-fits-all metabarcoding primer for all analyses, and there are numerous distinct primers designed to capture diverse taxonomic groups (Seymour et al., 2020). Certain primers that perform well with bulk sampling (e.g., insects blended together) will not perform as well with eDNA samples because the primers will sequence unwanted species in the sample, commonly protists or bacteria, that are more frequent (i.e., have more DNA available for sequencing). However, as the use of eDNA becomes more popular, new primers are being

produced that consider the type of environmental sample as well as the desired taxonomic diversity (Leese et al., 2021).

CONCLUSION

The region identified with the forward primer 477F has potential of being used as a metabarcoding region since the sequences are well conserved across species that can help in picking up multiple sxtA containing species with a few nucleotides variation between the species that can be targeted for designing a more species-specific probe. The primary difference required among the potential candidates is that they are to be constructed to have the variable region occur in either the middle of the primer, or at the 5’ or 3’ end to select for a more species-specific approach.

Future studies should focus on testing the primer probe sets identified for the 3’ end of the sxtA gene in cultured samples to test efficacy. Further testing should focus on testing the primer probe sets on environmental samples to assess the potential of one of the primer probe pairs to be developed into a species specific probe for early detection of saxitoxin-producing HABs. Also, attention is required in producing more complete genomic sequences of the entire gene cluster of Saxitoxin producing gene and a similar study to be repeated with more data.

As previously mentioned, using current methodologies of morphological analyses such as brightfield microscopy, many species look identical. It is imminent that we look for advanced molecular methodologies and develop novel probes as suggested for more accurate and reliable identification of toxic species that can help in early detection of toxic HABs.

LITERATURE CITED

Berry J, Gantar M, Perez M, Berry G, Noriega F. 2008. Cyanobacterial toxins as allelochemicals with potential applications as algaecides, herbicides and insecticides. Mar Drugs. 6(2):117–146.

Camacho FG, Rodríguez JG, Mirón AS, García MCC, Belarbi EH, Chisti Y, Grima EM. 2007. Biotechnological significance of toxic marine dinoflagellates. Biotechnol Adv. 25(2):176–194.

Christensen VG, Khan E. 2020. Freshwater neurotoxins and concerns for human, animal, and ecosystem health: A review of anatoxin-a and saxitoxin. Sci Total Environ. 736(139515):139515. doi:10.1016/j.scitotenv.2020.139515. https://www.sciencedirect.com/science/article/pii/S0048969720330321.

Cohen NJ, Deeds JR, Wong ES, Hanner RH, Yancy HF, White KD, Thompson TM, Wahl M, Pham TD, Guichard FM, et al. 2009. Public health response to puffer fish (Tetrodotoxin) poisoning from mislabeled product. J Food Prot. 72(4):810–817. doi:10.4315/0362-028x-72.4.810. [accessed 2022 Feb 15]. https://meridian.allenpress.com/jfp/article/72/4/810/173412/Public-Health-Response-to-Puffer-Fish-Tetrodotoxin.

Cristescu ME. 2019. Can environmental RNA revolutionize biodiversity science? Trends Ecol Evol. 34(8):694–697.

Cusick KD, Sayler GS. 2013. An overview on the marine neurotoxin, saxitoxin: genetics, molecular targets, methods of detection and ecological functions. Mar Drugs. 11(4):991–1018. doi:10.3390/md11040991. [accessed 2022 Feb 21]. https://www.ncbi.nlm.nih.gov/labs/pmc/articles/PMC3705384/.

Darling AE, Mau B, Perna NT. 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One. 5(6):e11147.0

Darling JA, Frederick RM. 2018. Nucleic acids-based tools for ballast water surveillance, monitoring, and research. J Sea Res. 133:43–52.

Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32(5):1792–1797.

Etheridge SM. 2010. Paralytic shellfish poisoning: seafood safety and human health perspectives. Toxicon. 56(2):108–122.

Fichez R, Jickells TD, Edmunds HM. 1992. Algal blooms in high turbidity, a result of the conflicting consequences of turbulence on nutrient cycling in a shallow water estuary. Estuar Coast Shelf Sci. 35(6):577–592.

Gessner BD, Bell P, Doucette GJ, Moczydlowski E, Poli MA, Van Dolah F, Hall S. 1997. Hypertension and identification of toxin in human urine and serum following a cluster of mussel-associated paralytic shellfish poisoning outbreaks. Toxicon. 35(5):711–722.

Gessner BD, Middaugh JP. 1995. Paralytic shellfish poisoning in Alaska: a 20-year retrospective analysis. Am J Epidemiol. 141(8):766–770.

Gill S, Murphy M, Clausen J, Richard D, Quilliam M, MacKinnon S, LaBlanc P, Mueller R, Pulido O. 2003. Neural injury biomarkers of novel shellfish toxins, spirolides: a pilot study using immunochemical and transcriptional analysis. Neurotoxicology. 24(4–5):593–604.

Hans W. Paerl, Rolland S. Fulton, Pia H. Moisander, and Julianne Dyble. 2001. Harmful Freshwater Algal Blooms, With an Emphasis on Cyanobacteria. The Scientific World

Journal. 1:38. doi:10.1100/tsw.2001.16. [accessed 2021 Oct 10]. https://www.hindawi.com/journals/tswj/2001/139109/.

Haselkorn R. 2009. Cyanobacteria. Current Biology. 19(7):R277-8.

Horner RA, Garrison DL, Plumley FG. 1997. Harmful algal blooms and red tide problems on the U.S. west coast. Limnol Oceanogr. 42(5part2):1076–1088.

Hurley W, Wolterstorff C, MacDonald R, Schultz D. 2014. Paralytic shellfish poisoning: a case series. West J Emerg Med. 15(4):378–381.

John D. Wehr and Robert G. Sheath. 2003. Freshwater algae of north America. Elsevier.

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, et al. 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 28(12):1647–1649. doi:10.1093/bioinformatics/bts199. https://www.ncbi.nlm.nih.gov/labs/pmc/articles/PMC3371832/.

Kellmann R, Mihali TK, Jeon YJ, Pickford R, Pomati F, Neilan BA. 2008. Biosynthetic intermediate analysis and functional homology reveal a saxitoxin gene cluster in cyanobacteria. Appl Environ Microbiol. 74(13):4044–4053. doi:10.1128/AEM.00353-08. [accessed 2022 Feb 10]. https://www.ncbi.nlm.nih.gov/labs/pmc/articles/PMC2446512/.

Liu Q, Zhang Y, Wu H, Liu F, Peng W, Zhang X, Chang F, Xie P, Zhang H. 2020. A review and perspective of eDNA application to eutrophication and HAB control in freshwater and marine ecosystems. Microorganisms. 8(3):417.

Marshall WS, Bryson SE, Darling P, Whitten C, Patrick M, Wilkie M, Wood CM, Buckland-Nicks J. 1997. NaCl transport and ultrastructure of opercular epithelium from a freshwater-adapted euryhaline teleost,Fundulus heteroclitus. J Exp Zool. 277(1):23–37.

Murray SA, Mihali TK, Neilan BA. 2011. Extraordinary conservation, gene loss, and positive selection in the evolution of an ancient neurotoxin. Mol Biol Evol. 28(3):1173–1182. doi:10.1093/molbev/msq295. [accessed 2022 Feb 10]. https://pubmed.ncbi.nlm.nih.gov/21076133/.

Murray SA, Wiese M, Stüken A, Brett S, Kellmann R, Hallegraeff G, Neilan BA. 2011. sxtA-based quantitative molecular assay to identify saxitoxin-producing harmful algal blooms in marine waters. Appl Environ Microbiol. 77(19):7050–7057.

National Oceanic and Atmospheric Administration U.S. Department of Commerce Enter Search Terms. 2016 Apr 27. What is a harmful algal bloom? Noaa.gov. [accessed 2021 Oct 10]. https://www.noaa.gov/what-is-harmful-algal-bloom.'

Ogram A, Sayler GS, Barkay T. 1987. The extraction and purification of microbial DNA from sediments. J Microbiol Methods. 7(2–3):57–66.

Orr RJS, Stüken A, Murray SA, Jakobsen KS. 2013. Evolution and distribution of saxitoxin biosynthesis in dinoflagellates. Mar Drugs. 11(8):2814–2828. doi:10.3390/md11082814. [accessed 2022 Feb 10]. https://www.ncbi.nlm.nih.gov/labs/pmc/articles/PMC3766867/.

Pilliod DS, Goldberg CS, Laramie MB, Waits LP. 2008. USGS fact sheet 2012–3146: Application of Environmental DNA for Inventory and Monitoring of aquatic species. Usgs.gov. [accessed 2021 Oct 15]. https://pubs.usgs.gov/fs/2012/3146/.

Pitcher GC, Figueiras FG, Hickey BM, Moita MT. 2010. The physical oceanography of upwelling systems and the development of harmful algal blooms. Prog Oceanogr. 85(1–2):5–32.

Rastogi RP, Madamwar D, Incharoensakdi A. 2015. Bloom dynamics of Cyanobacteria and their toxins: Environmental health impacts and mitigation strategies. Front Microbiol. 6:1254.

Rozen S, Skaletsky H. 2000. Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol. 132:365–386. doi:10.1385/1-59259-192-2:365. http://dx.doi.org/10.1385/1-59259-192-2:365.

Schantz EJ, Ghazarossian VE, Schnoes HK, Strong FM, Springer JP, Pezzanite JO, Clardy J. 1975. Structure of saxitoxin. J Am Chem Soc. 97(5):1238–1239. doi:10.1021/ja00838a045. http://dx.doi.org/10.1021/ja00838a045.

Schwartz S, Elnitski L, Li M, Weirauch M, Riemer C, Smit A, NISC Comparative Sequencing Program, Green ED, Hardison RC, Miller W. 2003. MultiPipMaker and supporting tools: Alignments and analysis of multiple genomic DNA sequences. Nucleic Acids Res. 31(13):3518–3524.

Schwarz EN. 2011. Molecular and morphological characterization of alexandrium species (dinophyceae) from the east coast, USA. University of North Carolina Wilmington.

Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, et al. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 7(1):539. doi:10.1038/msb.2011.75. https://pubmed.ncbi.nlm.nih.gov/21988835/.

Shimizu Y. 1996. Microalgal metabolites: a new perspective. Annu Rev Microbiol. 50(1):431–465. doi:10.1146/annurev.micro.50.1.431. [accessed 2022 Feb 10]. https://pubmed.ncbi.nlm.nih.gov/8905087/.

Soyer-Gobillard MO, Ausseil J, Géraud ML. 1996. Nuclear and cytoplasmic actin in dinoflagellates. Biol Cell. 87(1–2):17–35.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 30(9):1312–1313.

Stüken A, Orr RJS, Kellmann R, Murray SA, Neilan BA, Jakobsen KS. 2011. Discovery of nuclear-encoded genes for the neurotoxin saxitoxin in dinoflagellates. PLoS One. 6(5):e20096. doi:10.1371/journal.pone.0020096. [accessed 2022 Feb 15]. https://www.ncbi.nlm.nih.gov/labs/pmc/articles/PMC3097229/.

Seymour M, Edwards FK, Cosby BJ, Bista I, Scarlett PM, Brailsford FL, Glanville HC, de Bruyn M, Carvalho GR, Creer S. 2021. Environmental DNA provides higher resolution assessment of riverine biodiversity and ecosystem function via spatio-temporal nestedness and turnover partitioning. Commun Biol. 4(1):512. doi:10.1038/s42003-021-02031-2. [accessed 2022 Feb 24]. https://pubmed.ncbi.nlm.nih.gov/33941836/.

US Department of Commerce, National Oceanic, Atmospheric Administration. 2014. Why do harmful algal blooms occur? [accessed 2021 Oct 26]. https://oceanservice.noaa.gov/facts/why_habs.html.

US Department of Commerce, National Oceanic, Atmospheric Administration (NOAA). 2009. What are phytoplankton? [accessed 2021 Aug 26]. https://oceanservice.noaa.gov/facts/phyto.html.

Wiese M, D'Agostino PM, Mihali TK, Moffitt MC, Neilan BA. 2010. Neurotoxic alkaloids: saxitoxin and its analogs. Mar Drugs. 8(7):2185–2211. doi:10.3390/md8072185. [accessed 2022 Feb 10]. https://www.ncbi.nlm.nih.gov/labs/pmc/articles/PMC2920551/.

Zaiko A, Pochon X, Garcia-Vazquez E, Olenin S, Wood SA. 2018. Advantages and limitations of environmental DNA/RNA tools for marine biosecurity: Management and surveillance of non-indigenous species. Front Mar Sci. 5. doi:10.3389/fmars.2018.00322. http://dx.doi.org/10.3389/fmars.2018.00322.

Table 1. Organism's name and type used

Organism	Type of Organism
Alexandrium fundyense	dinoflagellates
Alexandrium tamarense	dinoflagellates
Alexandrium minutum	dinoflagellates
Pyrodinium bahamense	dinoflagellates
Alexandrium catenella	dinoflagellates
Gymnodinium catenatum	dinoflagellates
Alexandrium diversaporum	dinoflagellates
Alexandrium ostenfeldii	dinoflagellates
Alexandrium pacificum	dinoflagellates
Alexandrium pseudogonyaulax	dinoflagellates
Alexandrium fraterculus	dinoflagellates
Alexandrium affine	dinoflagellates
Alexandrium margalefii	dinoflagellates
Symbiodinium sp. CCMP2456	dinoflagellates
uncultured cyanobacterium	cyanobacteria
uncultured Dolichospermum sp.	cyanobacteria
Microseira wollei	cyanobacteria
Dolichospermum circinale AWQC131C	cyanobacteria
Cylindrospermopsis raciborskii T3	cyanobacteria
uncultured Anabaena sp.	Cyanobacteria
All other taxa	Others

Table 2. Accession number of the sequences obtained from NCBI's Nucleotide database

JF343240	JF343241	JF343242	JF343243	JF343244	JF343245
JF343246	JF343247	JF343248	JF343249	JF343250	JF343251
JF343252	JF343253	JF343254	JF343255	JF343256	JF343257
JF343258	JF343259	JF343260	JF343261	JF343262	JF343263
JF343264	JF343265	JF343266	JF343267	JF343268	JF343269
JF343270	JF343271	JF343272	JF343273	JF343274	JF343275
JF343276	JF343277	JF343278	JF343279	JF343280	JF343281
JF343282	JF343283	JF343284	JF343285	JF343287	JF343288
JF343289	JF343290	JF343291	JF343294	JF343295	JF343296
JF343297	JF343298	JF343299	JF343300	JF343301	JF343302
JF343303	JF343304	JF343305	JF343306	JF343307	JF343308
JF343309	JF343310	JF343311	JF343312	JF343313	JF343314
JF343315	JF343316	JF343317	JF343318	JF343319	JF343320
JF343321	JF343322	JF343323	JF343324	JF343325	JF343326
JF343327	JF343328	JF343329	JF343330	JF343331	JF343332
JF343333	JF343334	JF343335	JF343336	JF343337	JF343338
JF343339	JF343340	JF343341	JF343342	JF343343	JF343344
JF343345	JF343346	JF343347	JF343348	JF343349	JF343350
JF343351	JF343352	JF343353	JF343354	JF343355	JF343356
KC835398	KC835399	KC835400	KC835401	KC835402	KF985177
KF985178	KF985179	KF985180	KF985181	KF985182	KJ879194
KJ879195	KJ879196	KJ879197	KJ879198	KJ879199	KJ879200
KJ879201	KJ879202	KJ879203	KJ879204	KJ879205	KJ879206
KJ879207	KJ879208	KJ879209	KJ879210	KJ879211	KJ879212
KJ879213	KJ886938	KJ999785	KJ999786	KJ999787	KJ999788
KM100452	KM100453	KM100454	KM100455	KM104226	KM104227
KM104228	KM104229	KM104231	KM104232	KM104250	KM104251
KM104252	KM104253	KM104254	KM104255	KM104256	KM104257
KM104258	KM104259	KM104260	KM104261	KM104262	KM104263
KM104264	KM104265	KM104266	KM104267	KM104268	KM104269
KM104270	KM104271	KM104272	KM104273	KM104274	KM104275
KM104276	KM104277	KM104278	KM104281	KM104282	KM104283
KM104284	KM104285	KM104286	KM104287	KM104288	KM104289
KM104290	KM104291	KM104292	KM104293	KM104294	KM104295
KM104296	KM104297	KM104298	KM104299	KM104300	KM104301
KM104302	KM104303	KM104304	KM104305	KM104306	KM104307
KM104308	KM104309	KM104310	KM104311	KM104312	KM104313
KM104314	KM104315	KM104316	KM205606	KM205607	KM438016
KM438017	KM438018	KM438019	KM438020	KM438021	KM438022
KM438023	KM438024	KM438025	KM438026	KM438027	KY575968
KY575969	KY575970	KY575971	KY575972	KY575973	KY575974
LC547947	LC547948	LC547949	LC547950	LC547951	LC547952
LC547953	LC547954	LC547955	LC547956	LC547957	LC547958

LC547959	LC547960	MW248860	MW546902	MW546903	MW546904
MW546905	MW546906	MW546907	MW546908	MZ234664	MZ234665
MZ234666	MZ234667	MZ234668	MZ234669	MZ234670	MZ234671
MZ234672	MZ234673	MZ234674	MZ234675	MZ234676	MZ234677
MZ234678	MZ234679	MZ234680	MZ234681	MZ234682	MZ234683
MZ234684	MZ234685	MZ234686	MZ234687	MZ234688	MZ234689
MZ234690	MZ234691	MZ234692	MZ234693	MZ234694	MZ234695

Table 3. Complete list of Forward Primer sequences

Sl.No	5′ anchoring end sequence	Direction	Binding site(bp)	Name	GC-content
01	TGGACTACGCGGAGAACAAC	Forward	476-497	476F	55.0
02	GGACTACGCGGAGAACAACA	Forward	477-498	477F	55.0
03	TGTACCCGAGCGTTACCATG	Forward	49-68	49F	55.0
04	CATCTGGAAGTGCGTCCTGT	Forward	32-51	32F	55.0
05	GACATCACCTACGTGTGCGA	Forward	266-285	266F	55.0

Table 4. Percentage identity of 477F primer against the 206 sequences in the 3′ cluster

Sequence accession number	% identity with 477F
KC835402	100
KC835401	100
JF343328	100
JF343350	100
JF343310	100
JF343347	100
JF343316	100
KM438018	100
KM438017	100
JF343317	100
JF343312	100
JF343313	100
JF343314	100
JF343349	100
JF343325	100
JF343324	100
JF343315	100
JF343348	100
JF343326	100
JF343327	100
JF343311	100
JF343351	100
JF343265	100
KM438016	100
KM438019	100
KM438020	100
KM438023	100
MW546904	100
MW546903	100
MW546902	100
KM438022	100
MZ234686	100
MZ234694	100
MZ234682	100
MZ234685	100
MZ234667	100
MZ234675	100
MZ234688	100
MZ234672	100
MZ234681	100
MZ234690	100
MZ234666	100

MZ234692	100
MZ234693	100
MZ234677	100
MZ234679	100
MZ234676	100
MZ234678	100
MZ234665	100
MZ234683	100
MZ234664	100
KM104282	100
JF343343	100
JF343345	100
KM104267	100
KM104290	100
KM104268	100
KM100454	100
KJ886938	100
KM104269	100
KM104273	100
KM104274	100
KM104276	100
MW546908	100
JF343263	100
JF343333	100
JF343332	100
JF343264	100
JF343335	100
JF343334	100
KF985181	100
JF343262	100
JF343336	100
KJ999788	100
KM104275	100
KM104272	100
JF343340	100
LC547955	100
JF343338	100
KM104257	100
KM104259	100
KM104262	100
LC547953	100
KJ879198	100
KM100453	100
KM104284	100
KJ879203	100

KJ879197	100
KM100452	100
KM104266	100
JF343344	100
KJ879201	100
KF985179	100
KM104250	100
KJ879212	100
KM104258	100
LC547956	100
JF343329	100
JF343260	100
KJ879196	100
KM100455	100
KM104255	100
MW546907	100
JF343342	100
JF343356	100
JF343261	100
JF343259	100
KF985178	100
KF985177	100
LC547957	100
LC547954	100
KJ879208	100
MW546906	100
MW546905	100
KM104281	100
KJ879213	100
KJ999785	100
KJ999787	100
KM104277	100
KM104263	100
KM104287	100
KJ879211	100
KJ879205	100
JF343341	100
LC547951	100
JF343353	100
LC547947	100
LC547958	100
LC547959	100
LC547950	100
LC547948	100
KM104260	100

LC547960	100
JF343339	100
KM104289	96.25
KJ879210	96.25
KJ879199	96.25
KJ879195	96.25
KJ879204	96.25
KF985180	96.25
JF343323	95
JF343319	95
JF343322	95
JF343320	95
JF343318	95
JF343321	95
KM438024	95
KM438025	95
KM438021	95
MZ234687	95
MZ234673	95
MZ234674	95
MZ234668	95
MZ234670	95
MZ234671	95
MZ234680	95
MZ234691	95
MZ234684	95
KM104300	95
KJ999786	95
KM104261	95
KM104278	95
KJ879207	95
KM104252	95
KM104253	95
KM104256	95
KM104285	95
JF343354	95
LC547949	95
KM104270	95
KM104271	95
KJ879209	95
JF343337	95
KM104296	95
KM104293	95
KM104254	95
JF343346	95

KF985182	95
KJ879206	95
LC547952	95
JF343355	95
KM104299	95
KJ879194	95
KM104294	95
KM104298	95
KM104291	95
KM104297	95
KM104292	95
KM104295	95
KM104251	95
KJ879202	95
KJ879200	90
KM104301	90
JF343330	90
JF343331	90
MZ234689	86.25
JF343352	85
KM104286	85
KM104265	85
KM104264	85
MZ234695	85
MZ234669	85
KM104283	70
Consensus	50
KM104288	0

Table 5. Percentage identity of 476F primer against the 206 sequences in the 3′ cluster

Sequence accession number	% identity with 477F
KC835402	100
KC835401	100
JF343328	100
JF343350	100
JF343310	100
JF343347	100
JF343316	100
KM438018	100
KM438017	100
JF343317	100
JF343312	100
JF343313	100
JF343314	100
JF343349	100
JF343325	100
JF343324	100
JF343315	100
JF343348	100
JF343326	100
JF343327	100
JF343311	100
JF343351	100
JF343265	100
KM438016	100
KM438019	100
KM438020	100
KM438023	100
MW546904	100
MW546903	100
MW546902	100
KM438022	100
MZ234686	100
MZ234694	100
MZ234682	100
MZ234685	100
MZ234667	100
MZ234675	100
MZ234688	100
MZ234672	100
MZ234681	100
MZ234690	100
MZ234666	100
MZ234692	100

MZ234693	100
MZ234677	100
MZ234679	100
MZ234676	100
MZ234678	100
MZ234665	100
MZ234683	100
MZ234664	100
KM104282	100
JF343343	100
JF343345	100
KM104267	100
KM104290	100
KM104268	100
KM100454	100
KJ886938	100
KM104269	100
KM104273	100
KM104274	100
KM104276	100
MW546908	100
JF343263	100
JF343333	100
JF343332	100
JF343264	100
JF343335	100
JF343334	100
KF985181	100
JF343262	100
JF343336	100
KJ999788	100
KM104275	100
KM104272	100
JF343340	100
LC547955	100
JF343338	100
KM104257	100
KM104259	100
KM104262	100
LC547953	100
KJ879198	100
KM100453	100
KM104284	100
KJ879203	100
KJ879197	100
KM100452	100

KM104266	100
JF343344	100
KJ879201	100
KF985179	100
KM104250	100
KJ879212	100
KM104258	100
LC547956	100
JF343329	100
JF343260	100
KJ879196	100
KM100455	100
KM104255	100
MW546907	100
JF343342	100
JF343356	100
JF343261	100
JF343259	100
KF985178	100
KF985177	100
LC547957	100
LC547954	100
KJ879208	100
MW546906	100
MW546905	100
KM104281	100
KJ879213	100
KJ999785	100
KJ999787	100
KM104277	100
KM104263	100
KM104287	100
KJ879211	100
KJ879205	100
JF343341	100
LC547951	100
JF343353	100
LC547947	100
LC547958	100
LC547959	100
LC547950	100
LC547948	100
KM104260	100
LC547960	100
JF343339	100
KM104289	96.25

KJ879210	96.25
KJ879199	96.25
KJ879195	96.25
KJ879204	96.25
KF985180	96.25
JF343323	95
JF343319	95
JF343322	95
JF343320	95
JF343318	95
JF343321	95
KM438024	95
KM438025	95
KM438021	95
MZ234687	95
MZ234673	95
MZ234674	95
MZ234668	95
MZ234670	95
MZ234671	95
MZ234680	95
MZ234691	95
MZ234684	95
KM104300	95
KJ999786	95
KM104261	95
KM104278	95
KJ879207	95
KM104252	95
KM104253	95
KM104256	95
KM104285	95
JF343354	95
LC547949	95
KM104270	95
KM104271	95
KJ879209	95
JF343337	95
KM104296	95
KM104293	95
KM104254	95
JF343346	95
KF985182	95
KJ879206	95
LC547952	95
JF343355	95

KM104299	95
KJ879194	95
KM104294	95
KM104298	95
KM104291	95
KM104297	95
KM104292	95
KM104295	95
KM104251	95
KJ879202	95
JF343330	95
JF343331	95
MZ234695	90
MZ234669	90
KJ879200	90
KM104301	90
MZ234689	86.25
JF343352	85
KM104286	85
KM104265	85
KM104264	85
KM104283	70
Consensus	45
KM104288	0