



Exploring the Microbiota of the Guarapiranga Water Reservoir With Long-Read Sequencing Technology

Douglas M. M. Soares¹, Samir V. F. Atum^{1,2}, Eteivino J. H. Bechara¹, João C. Setubal², Cassius V. Stevani¹ and Renato S. Freire^{1*}

¹ Departamento de Química Fundamental, Instituto de Química, Universidade de São Paulo, São Paulo, Brazil,

² Departamento de Bioquímica, Instituto de Química, Universidade de São Paulo, São Paulo, Brazil

Keywords: *Caudovirales*, environmental metagenomics, freshwater, nanopore, *proteobacteria*, water supply

INTRODUCTION

The water reservoir of Guarapiranga supplies drinking water to more than 3.7 million people in the São Paulo (Brazil) metropolitan area, one of the most densely populated and industrialized regions worldwide. This reservoir is also used for recreational purposes such as fishing, bathing and aquatic sports. As some densely populated areas abut its shore, the water quality may be affected by diffuse pollution discharges. Eutrophication is a recurrent problem for maintaining the quality of water in this reservoir, mainly due to the risk of toxin-producing cyanobacteria blooms. Since the early 1970s, the Guarapiranga reservoir has remained eutrophic (Mozeto et al., 2001). Copper sulfate has been used as one of the strategies to control bacterial blooms (Beyruth, 2000). Nevertheless, the occurrence of worrisome amounts of cyanobacteria in Guarapiranga is still reported in recent studies (Alcantara et al., 2021).

Many actinomycetes and cyanobacteria are well known for their ability to produce diverse secondary metabolites including toxins (such as microcystin, saxitoxin and cylindrospermopsin) and taste-and-odor compounds [such as geosmin and 2-methylisoborneol (2-MIB)] (Sivonen and Jones, 1999; Graham et al., 2010). Cyanotoxins may be classified based on toxic mechanism, in particular hepatotoxins (e.g., cylindrospermopsins and microcystins) and neurotoxins (e.g., anatoxins and saxitoxins). Many cyanotoxins have multiple variants with a range of toxicities (Welker and von Döhren, 2006). Several studies have reported the negative health and ecological effects of cyanotoxins (Carmichael, 1992; Funari and Testai, 2008). Therefore, Brazil and several other countries have set national standards for cyanotoxins in drinking and recreational waters (Hudnell, 2008). Microcystins are the most common cyanotoxins in Brazilian aquatic ecosystems (Genuário et al., 2016). Unlike toxins, the taste-and-odor compounds are not actually toxic and, usually, there are no regulations for the presence of these compounds in drinking water. However, taste-and-odor compounds are of concern to the consuming public, and they are the primary criteria of drinking water safety considered by consumers. Moreover, the co-occurrence of cyanotoxins and taste-and-odor compounds have been studied in surface waters worldwide. For instance, microcystin co-occurred with geosmin in 87% and 2-MIB in 39% of the cyanobacterial blooms sampled from 23 Midwestern United States lakes (Graham et al., 2010).

Due to the potential impacts of algal and bacterial blooms to water quality, information and identification of public health-relevant microbial species in drinking water reservoirs is important. The metagenomic approach overcomes microbial culturing limitations, providing taxonomic and functional diversity information of the sampled environment. Despite the importance of the water reservoir of Guarapiranga, literature on its microbial diversity and

OPEN ACCESS

Edited by:

Gustavo Fonseca,
Federal University of São Paulo, Brazil

Reviewed by:

Gustavo Bueno Gregoracci,
Federal University of São Paulo, Brazil

Fabiana S. Paula,

University of São Paulo, Brazil

Yong Yu,

Polar Research Institute of
China, China

*Correspondence:

Renato S. Freire
rsfreire@iq.usp.br

Specialty section:

This article was submitted to
Marine Evolutionary Biology,
Biogeography and Species Diversity,
a section of the journal
Frontiers in Marine Science

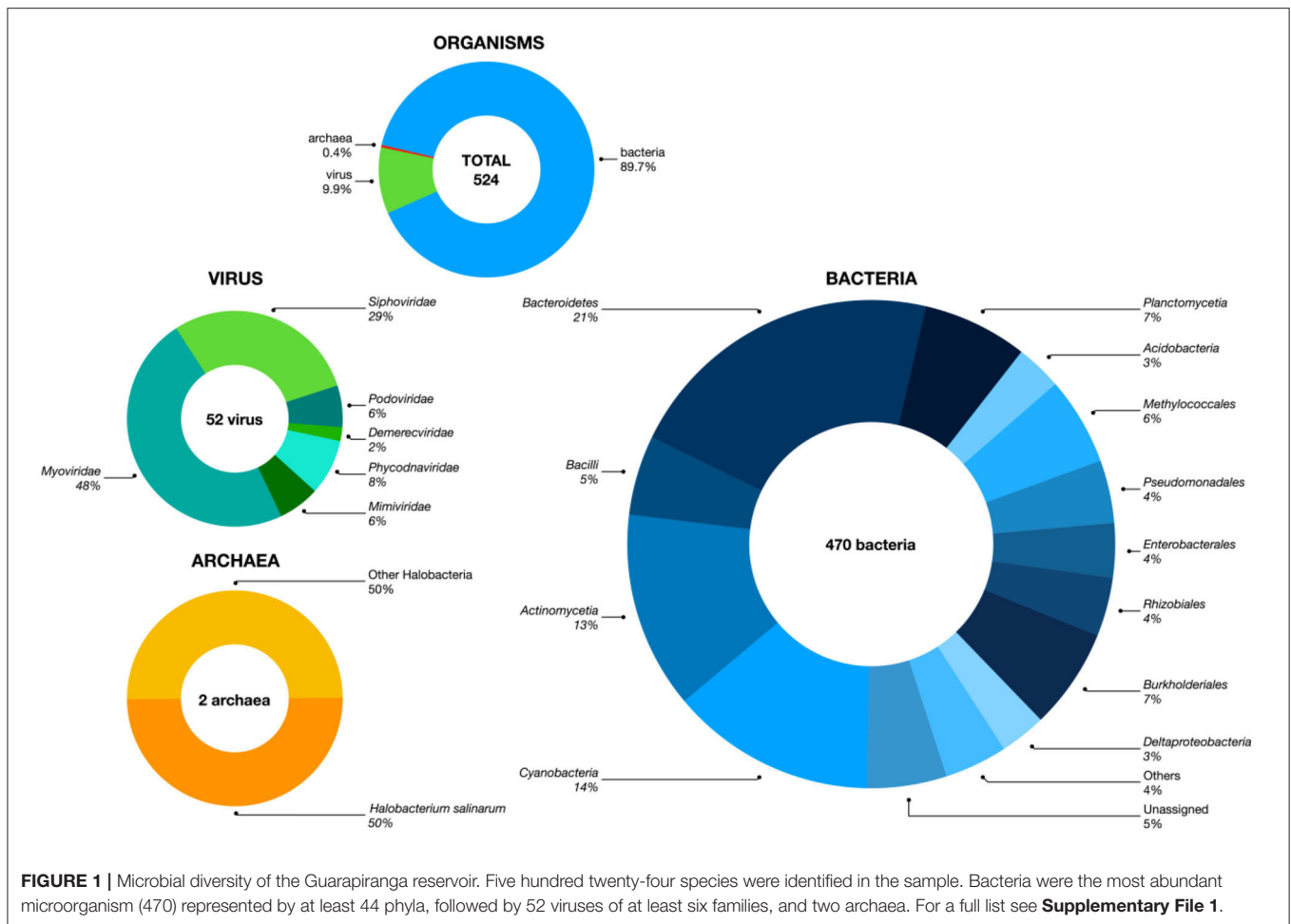
Received: 08 October 2021

Accepted: 09 November 2021

Published: 03 December 2021

Citation:

Soares DMM, Atum SVF,
Bechara EJH, Setubal JC, Stevani CV
and Freire RS (2021) Exploring the
Microbiota of the Guarapiranga Water
Reservoir With Long-Read
Sequencing Technology.
Front. Mar. Sci. 8:791101.
doi: 10.3389/fmars.2021.791101



pollution impact is scarce (Fontana et al., 2014; Alcántara et al., 2021; Pierangeli et al., 2021). Herein we report the first metagenomic analysis of surface water from the Guarapiranga reservoir. Microbial diversity and populational dynamics of this area can help assess the impact of pollution, eutrophication risk, and proliferation of toxin-bearing cyanobacteria, which can be useful for developing more efficient water treatment strategies to assure high standards of public health.

METHODS AND DATA ANALYSIS

Surface water samples were collected from Guarapiranga reservoir near to the water abstraction point used by the water treatment plant at coordinates 23°40'23"S, 46°43'12"W. Samples were collected in December 2020 in the euphotic zone (0.5 m depth), following the methodology described by the Brazilian National Guide of Sampling Collection and Preservation (CETESB/ANA., 2011), kept at 4°C for transportation and processed within 24 h.

A total volume of 900 mL of pooled sample water was filtered using 0.22- μ m pore bottle-top vacuum filter system (KASVI). After total filtration, the filter membrane was removed using two sets of sterile forceps and inserted into

the PowerWater DNA bead tube, with the top side facing inward. Total genomic DNA was extracted with DNeasy® PowerWater® kit (QIAGEN), followed by an additional cleaning step using DNeasy columns (QIAGEN) for an efficient removal of contaminants. DNA concentration and purity were evaluated by both NanoPhotometer NP80® (IMPLEN) and Qubit™ fluorometer (Life Technologies).

The sequencing library was prepared from 450 ng of genomic DNA with the Rapid DNA Sequencing kit (SQK-RAD004) from Oxford Nanopore Technologies (ONT). DNA sequencing was performed using the minION device (ONT) with the FLO-MIN106 flow cell, and the quality of sequencing monitored in real time on the MinKNOW interface (ONT). Computational processing was as follows (default parameters were used unless otherwise noted): Basecalling was performed by Guppy v.4.4.1 with the high quality r9.4.1_450bps_hac model. Reads were assembled with metaFlye v.2.8.3 with the `-meta` and `-plasmids` flags (Kolmogorov et al., 2020). Contamination filtering was done by comparing contigs with the NCBI UniVec database¹. Contigs that passed this check were classified with kaiju v.1.7.4 (Menzel et al., 2016), with the RefSeq (2021-02-26) database downloaded

¹<https://www.ncbi.nlm.nih.gov/tools/vecscreen/univec/>

from the Kaiju Webserver². Functional annotation was provided by the IMG/JGI annotation pipeline (Chen et al., 2021), and the assembly is available from JGI GOLD (Mukherjee et al., 2020). Search for secondary metabolites in the assembly was performed with antiSMASH version 6 (Blin et al., 2021).

A total of 512,095 reads were obtained, with average length of 2,282 bp. The assembly has 655 contigs and a total of 18,457,687 bp, with a N50 of 41,334 bp. Bacteria correspond to 89.7% of the assembled metagenome (**Figure 1, Supplementary File 1**), and the most abundant bacterial phyla identified were *Proteobacteria* (35%); *Terrabacteria* (28%); Fibrobacteres, *Chlorobi*, and *Bacteroidetes* (FCB group, 17%); *Planctomycetes*, *Verrucomicrobia*, and *Chlamydiae* (PVC group, 8%); and *Acidobacteria* (7%). Viruses represent 9.9% of the assembly, 85% of which are *Caudovirales bacteriophages*, and 15% are protist and algae-infecting *Megaviricetes*. Only 0.4% of the metagenome (two species) was identified as archaea, genus *Halobacteria*.

Due to the importance of cyanotoxins and taste-and-odor compounds to drinking water quality, the presence of genes related to biosynthesis of these substances were verified by using a manually-curated dataset of 3,265 sequences of genes from the NCBI gene database (**Supplementary File 2**). We then translated and compared these sequences with the assembled metagenome with tBLASTn. None of the 2,102 microcystin, 505 geosmin, 450 saxitoxin, 146 cylindrospermopsin, and 62 2-MIB biosynthesis genes were found in the Guarapiranga metagenome, even though we recovered genomic fragments from 49 cyanobacteria species. We hypothesize that organisms that express these compounds are either absent from the samples or present in such low concentrations that our sampling was not able to detect them. Moreover, it is known that massive amounts of copper sulfate and other chemicals have been used since 1970's by the São Paulo State agency responsible for basic sanitation (SABESP) for the control of algae and cyanobacteria (Leal et al., 2018).

The metagenomic analysis is the first step toward the identification of key organisms responsible to produce toxins and taste-and-odor compounds. The presence of these microorganisms is not a *sine qua non* condition for the release of such compounds in the water body, which surely depend on the expression of specific enzymes involved in the biosynthetic pathways. Nevertheless, the detection of these microorganisms is a precocious indication of a potential production of toxins and taste-and-odor compounds, which can warn water supplier companies of an imminent danger. Hence, metagenomics can contribute to water quality control management.

From the functional annotation results we highlight photosynthesis and carbon fixation genes like transketolases, phage-related genes like integrases and lysozymes; and transmembrane transport (**Supplementary File 3**). The majority of transmembrane transport genes are related to the ABC multidrug transport system or the AcrAB-TolC multidrug efflux pump. Both are known to be related to antibiotic resistance

(Abdi et al., 2020). Using AntiSMASH, we identified gene clusters related to the synthesis of proteusin, terpenes, arylpolene, and RRE-containing secondary metabolites.

Among identified species (**Supplementary File 4**) we highlight the following: *Aquirufa nivalisilvae* (contig 491, 251,358 bp), which belongs to a relatively new (2019) genus of widespread freshwater bacteria (Pitt et al., 2019); Candidatus *Nitrosacidococcus tergens* (contig 572, 144,006 bp), a species originally isolated from the biofilter unit of a pig farm in the Netherlands, able to grow and oxidize ammonia at pH 2.5 (Picone et al., 2021); and *Frigoriglobus tundricola* (contig 636, 180,833 bp), a species originally isolated from tundra wetland in Russia, and which is a cellulolytic planctomycete (Kulichevskaya et al., 2020).

DATA AVAILABILITY STATEMENT

The raw reads obtained in this project have been deposited in the NCBI Short Read Archive under the accession number PRJNA758495. The assembly and annotation are available under the JGI GOLD project ID Gp0563454.

AUTHOR CONTRIBUTIONS

DS: sample preparation and sequencing. SA: base calling, binning, assembling, annotations, and gene analysis. EB, JS, CS, and RF: experimental design, sampling, manuscript preparation, and miscellaneous stuff. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, 2017/22501-2) and CNPq. DS and SA received postdoctoral and doctoral scholarships from FAPESP (2019/12605-0) and the Brazilian Federal Agency for Support and Evaluation of Graduate Education (CAPES, 88887.514039/2020-00), respectively. JS and EB were funded in part by a CNPq Senior Researcher Fellowship.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmars.2021.791101/full#supplementary-material>

Supplementary File 1 | Relative abundances of microbial taxa in water sample collected from Guarapiranga reservoir (São Paulo, Brazil).

Supplementary File 2 | Curated dataset of taste-and-odor compounds (2-MIB and geosmin) cyanotoxins (cylindrospermopsin, microcystin and saxitoxin) genes.

Supplementary File 3 | Count of functional categories present in the assembly evaluated with different databases.

Supplementary File 4 | List of the contigs larger than 100kb, their sizes and their respective classifications given by Kaiju.

²<http://kaiju.binf.ku.dk/server>

REFERENCES

- Abdi, S. N., Ghotaslou, R., Ganbarov, K., Mobed, A., Tanomand, A., Yousefi, M., et al. (2020). *Acinetobacter baumannii* efflux pumps and antibiotic resistance. *Infect. Drug Resist.* 13, 423–434. doi: 10.2147/IDR.S228089
- Alcantara, E., Coimbra, K., Ogashawara, I., Rodrigues, T., Mantovani, J., Rotta, L. H., et al. (2021). A satellite-based investigation into the algae bloom variability in large water supply urban reservoirs during COVID-19 lockdown. *Remote Sens. Appl. Soc. Environ.* 23, 100555. doi: 10.1016/j.rsase.2021.100555
- Beyruth, Z. (2000). Periodic disturbances, trophic gradient and phytoplankton characteristics related to cyanobacterial growth in Guarapiranga Reservoir, São Paulo State, Brazil. *Hydrobiologia*. 424, 51–65. doi: 10.1023/A:1003944726346
- Blin, K., Shaw, S., Kloosterman, A. M., Charlop-Powers, Z., van Wezel, G. P., Medema, M. H., et al. (2021). antiSMASH 6.0: improving cluster detection and comparison capabilities. *Nucleic Acids Res.* 49, 0305–1048. doi: 10.1093/nar/gkab335
- Carmichael, W. W. (1992). Cyanobacteria secondary metabolites—the cyanotoxins. *J. Appl. Microbiol.* 72, 445–459. doi: 10.1111/j.1365-2672.1992.tb01858.x
- CETESB/ANA. (2011). Guia nacional de coleta e preservação de amostras: água, sedimento, comunidades aquáticas e efluentes líquidas São Paulo: CETESB. Brasília, DF. Available online at: <https://arquivos.ana.gov.br/institucional/sge/CEDOC/Catalogo/2012/GuiaNacionalDeColeta.pdf> (accessed August 12, 2021).
- Chen, I.-M. A., Chu, K., Palaniappan, K., Ratner, A., Huang, J., Huntemann, M., et al. (2021). The IMG/M data management and analysis system v.6.0: new tools and advanced capabilities. *Nucleic Acids Res.* 49, D751–D763. doi: 10.1093/nar/gkaa939
- Fontana, L., Albuquerque, A. L. S., Brenner, M., Bonotto, D. M., Sabaris, T. P. P., Pires, M. A. F., et al. (2014). The eutrophication history of a tropical water supply reservoir in Brazil. *J. Paleolimnol.* 51, 29–43. doi: 10.1007/s10933-013-9753-3
- Funari, E., and Testai, E. (2008). Human health risk assessment related to cyanotoxins exposure. *Crit. Rev. Toxicol.* 38, 97–125. doi: 10.1080/10408440701749454
- Genuário, D. B., Lorenzi, A. S., Aguiaro, L. F., Isaac, R. L., Azevedo, M. T. P., Cantúcio Neto, R., et al. (2016). Cyanobacterial community and microcystin production in a recreational reservoir with constant Microcystis blooms. *Hydrobiologia*. 779, 105–125. doi: 10.1007/s10750-016-2802-y
- Graham, J. L., Loftin, K. A., Meyer, M. T., and Ziegler, A. C. (2010). Cyanotoxin mixtures and taste-and-odor compounds in cyanobacterial blooms from the midwestern United States. *Environ. Sci. Technol.* 44, 7361–7368. doi: 10.1021/es1008938
- Hudnell, H. K. Cyanobacterial Harmful Algal Blooms: State of the Science and Research Needs. New York: Springer. (2008).
- Kolmogorov, M., Bickhart, D. M., Behsaz, B., Gurevich, A., Rayko, M., Shin, S. B., et al. (2020). MetaFlye: scalable long-read metagenome assembly using repeat graphs. *Nat. Methods*. 17, 1103–1110. doi: 10.1038/s41592-020-00971-x
- Kulichevskaya, I. S., Ivanova, A. A., Naumoff, D. G., Beletsky, A. V., Rijpstra, W. I. C., Damsté, J. S. S., et al. (2020). *Frigoriglobus tundricola* gen. nov., sp. nov., a psychrotolerant cellulolytic planctomycete of the family Gemmataceae from a littoral tundra wetland. *Syst. Appl. Microbiol.* 43, 126129. doi: 10.1016/j.syapm.2020.126129
- Leal, P. R., Moschini-Carlos, V., López-Doval, J. C., Cintra, J. P., Yamamoto, J. K., Bitencourt, M. D., et al. (2018). Impact of copper sulfate application at an urban Brazilian reservoir: A geostatistical and ecotoxicological approach. *Sci. Total Environ.* 618, 621–634. doi: 10.1016/j.scitotenv.2017.07.095
- Menzel, P., Ng, K., and Krogh, A. (2016). Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat. Commun.* 7, 11257. doi: 10.1038/ncomms11257
- Mozeto, A. A., Silvério, P. F., and Soares, A. (2001). Estimates of benthic fluxes of nutrients across the sediment–water interface (Guarapiranga reservoir, São Paulo, Brazil). *Sci. Total Environ.* 266, 135–142. doi: 10.1016/S0048-9697(00)00726-9
- Mukherjee, S., Stamatis, D., Bertsch, J., Ovchinnikova, G., Sundaramurthi, J. C., Lee, J., et al. (2020). Genomes onLine database (GOLD) v.8: overview and updates. *Nucleic Acids Res.* 49, D723–D733. doi: 10.1093/nar/gkaa983
- Picone, N., Pol, A., Mesman, R., van Kessel, M. A. H. J., Cremers, G., van Gelder, A. H., et al. (2021). Ammonia oxidation at pH 2.5 by a new gammaproteobacterial ammonia-oxidizing bacterium. *ISME J.* 15, 1150–1164. doi: 10.1038/s41396-020-00840-7
- Pierangeli, G. M. F., Domingues, M. R., Jesus, T. A., Coelho, L. H. G., Hanisch, W. S., Pompêo, M. L. M., et al. (2021). Higher abundance of sediment methanogens and methanotrophs do not predict the atmospheric methane and carbon dioxide flows in eutrophic tropical freshwater reservoirs. *Front. Microbiol.* 17, 647921. doi: 10.3389/fmicb.2021.647921
- Pitt, A., Schmidt, J., Kol, U., and Hahn, M. W. (2019). *Aquirufa antheringensis* gen. nov., sp. nov. and *Aquirufa nivalisilvae* sp. nov., representing a new genus of widespread freshwater bacteria. *Int. J. Syst. Evol. Microbiol.* 69, 2739–2749. doi: 10.1099/ijsem.0.003554
- Sivonen, K., and Jones, G. (1999). “Cyanobacterial toxins,” in *Toxic Cyanobacteria in Water: A Guide to Their Public Health Consequences, Monitoring and Management*, eds I. Chorus and J. Bartram (Bury St Edmunds, Suffolk: St Edmundsbury Press).
- Welker, M., and von Döhren, H. (2006). Cyanobacterial peptides: nature’s own combinational biosynthesis. *FEMS Microbiol. Rev.* 30, 530–563. doi: 10.1111/j.1574-6976.2006.00022.x

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

The reviewer FP declared a shared affiliation, though no other collaboration, with the authors.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Soares, Atum, Bechara, Setubal, Stevani and Freire. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.