

Metagenomics Analysis of Microorganisms in Freshwater Lakes of the Amazon Basin

Danyelle Toyama,^a Luciano Takeshi Kishi,^a Célio Dias Santos-Júnior,^a Andrea Soares-Costa,^a Tereza Cristina Souza de Oliveira,^b Fernando Pellon de Miranda,^c Flávio Henrique-Silva^a

Department of Genetics and Evolution, Center for Biological and Health Sciences, Federal University of São Carlos, São Carlos, São Paulo, Brazil^a; Department of Chemistry, Federal University of Amazonas, Manaus, Amazonas, Brazil^b; Petróleo Brasileiro S.A. (Petrobras), Centro de Pesquisas e Desenvolvimento Leopoldo Américo Miguez de Mello, Rio de Janeiro, Brasil^c

D.T. and L.T.K. contributed equally to this work.

The Amazon Basin is the largest hydrographic basin on the planet, and the dynamics of its aquatic microorganisms strongly impact global biogeochemical cycles. However, it remains poorly studied. This metagenome project was performed to obtain a snapshot of prokaryotic microbiota from four important lakes in the Amazon Basin.

Received 26 October 2016 Accepted 28 October 2016 Published 22 December 2016

Citation Toyama D, Kishi LT, Santos-Júnior CD, Soares-Costa A, de Oliveira TCS, de Miranda FP, Henrique-Silva F. 2016. Metagenomics analysis of microorganisms in freshwater lakes of the Amazon Basin. *Genome Announc* 4(6):e01440-16. doi:10.1128/genomeA.01440-16.

Copyright © 2016 Toyama et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Flávio Henrique-Silva, dfhs@ufscar.br.

The Solimões–Amazon River drains the Andes Mountains by crossing a humid and flat region to the Atlantic Ocean. This complex fluvial system carries 1,240 megatons of sediments per year and reworks approximately 3,200 megatons of floodplain sediments per year (1). Its annual discharge of 6.3 trillion m³ represents approximately 16% of all freshwater released into the world's oceans (2).

The Amazon region includes a large number of rivers and lakes, as well as the world's largest rainforest characterized by enormous biodiversity. However, the microbiome of this habitat has been poorly studied. In Brazil, the first freshwater metagenomics study performed in the Solimões River revealed a diverse community with a vast array of potential metabolisms in the largest hydrographic basin on the planet (3).

Freshwater environments and their microbial communities configure the basis of the food web and are the major agents of biogeochemical cycling (4). These environments differ from soil, marine, and freshwater environments in their composition of microbial communities. Temperature, pH, hydrological retention time, and trophic state can determine the composition of microbial communities of lakes (5, 6). Some lakes may receive water from nearby rivers in the flood season, which can change the microbial community of these environments. In the Amazon region, these lakes receive material from the Solimões River, for example. In this work, we performed metagenomics high-throughput sequencing to understand the ecology of microbial communities from lakes adjacent to the Solimões River.

We collected water samples in September 2008 (dry season) from Lake Poraquê (03°57'36.36"S; 63°09'48.17"W), Lake Preto (03°21'12.46"S; 60°37'31.30"W), Manacapuru Great Lake (03°15'50.96"S; 60°41'19.76"W), and Lake Ananá (03°53'11.83"S; 61°40'36.75"W). Sampling and total DNA extraction were conducted using a methodology described elsewhere (3). We sequenced metagenomic DNA with paired-end 2 × 100-bp reads on

an Illumina HiSeq 2500 platform (Illumina, Inc., San Diego, CA, USA). Reads were then subjected to quality trimming using the NGS QC toolkit version 2.3.3 software (7) by discarding low-quality data (<Q20). A total of 334 million sequences for Lake Preto, 349 million sequences for Manacapuru Great Lake, 376 million sequences for Lake Ananá, and 165 million sequences for Lake Poraquê were submitted for annotation and classification using the MG-RAST server (8).

Taxonomic assignment revealed that metagenomes contained 89 to 98% (of assigned reads) bacteria, 0.3 to 1.2% archaea, and 0.5 to 2.5% viruses. The major observed bacterial phyla included *Proteobacteria* (abundance percentages: 36.1 to 59.5%), *Cyanobacteria* (1.8 to 32.6%), *Actinobacteria* (12.7 to 28.2%), *Planctomycetes* (0.5 to 3.6%), *Bacteroidetes* (1.9 to 3.5%) and *Firmicutes* (2.3 to 5.5%). Functional annotations revealed a considerable fraction of reads associated with respiration (3.3 to 3.8%), photosynthesis (0.2 to 1.1%), and phages, prophages, transposable elements, and plasmids (1.6 to 3.6%). This metagenome project provides valuable information for future studies about the metabolic potential of microbial communities present in lakes in this region.

Accession number(s). The sequences obtained in this project have been deposited in the NCBI Short Read Archive under the following GenBank accession numbers: [SRX659579](https://www.ncbi.nlm.nih.gov/submit/SLR000000000/) (Lake Poraquê), [SRX659576](https://www.ncbi.nlm.nih.gov/submit/SLR000000000/) (Lake Preto), [SRX659577](https://www.ncbi.nlm.nih.gov/submit/SLR000000000/) (Manacapuru Great Lake), and [SRX659578](https://www.ncbi.nlm.nih.gov/submit/SLR000000000/) (Lake Ananá).

FUNDING INFORMATION

This study was supported by Petróleo Brasileiro S.A. (Petrobras) as part of a research agreement (no. 0050.0081178.13.9) with the Federal University of São Carlos, SP, Brazil, within the context of the Geochemistry Thematic Network. F.H.-S. is a recipient of a Research Productivity Scholarship from the National Council for Scientific and Technological Development (CNPq; no. 311745/2013-0). L.T.K. was the recipient of a post-doctoral scholarship and D.T. was the recipient of a doctoral scholarship from the Brazilian Federal Agency for Support and Evaluation of

Graduate Education (CAPES). C.D.S.J. is a recipient of a doctoral scholarship from the National Council for Scientific and Technological Development (CNPq).

REFERENCES

1. Mertes LAK, Dunne T, Martinelli LA. 1996. Channel-floodplain geomorphology along the Solimões-Amazon River, Brazil. *Geol Soc Am Bull* 108:1089–1107. [http://dx.doi.org/10.1130/0016-7606\(1996\)108<1089:CFGATS>2.3.CO;2](http://dx.doi.org/10.1130/0016-7606(1996)108<1089:CFGATS>2.3.CO;2).
2. Oltman RE. 1968. Reconnaissance investigations of the discharge and water quality of the Amazon River. Geological Survey Circular 552. U.S. Department of the Interior, Washington, DC.
3. Ghai R, Rodriguez-Valera F, McMahon KD, Toyama D, Rinke R, Cristina Souza de Oliveira T, Wagner Garcia J, Pellon de Miranda F, Henrique-Silva F. 2011. Metagenomics of the water column in the pristine upper course of the Amazon River. *PLoS One* 6:e23785. <http://dx.doi.org/10.1371/journal.pone.0023785>.
4. Percent SF, Frischer ME, Vescio PA, Duffy EB, Milano V, McLellan M, Stevens BM, Boylen CW, Nierzwicki-Bauer SA. 2008. Bacterial community structure of acid-impacted lakes: what controls diversity? *Appl Environ Microbiol* 74:1856–1868. <http://dx.doi.org/10.1128/AEM.01719-07>.
5. Newton RJ, Jones SE, Helmus MR, McMahon KD. 2007. Phylogenetic ecology of the freshwater *Actinobacteria* acI lineage. *Appl Environ Microbiol* 73:7169–7176. <http://dx.doi.org/10.1128/AEM.00794-07>.
6. Lindström ES, Kamst-Van Agterveld MP, Zwart G. 2005. Distribution of typical freshwater bacterial groups is associated with pH, temperature, and lake water retention time. *Appl Environ Microbiol* 71:8201–8206. <http://dx.doi.org/10.1128/AEM.71.12.8201-8206.2005>.
7. Patel RK, Jain M. 2012. NGS QC toolkit: a toolkit for quality control of next generation sequencing data. *PLoS One* 7:e30619. <http://dx.doi.org/10.1371/journal.pone.0030619>.
8. Meyer F, Paarmann D, D'Souza M, Olson R, Glass EM, Kubal M, Paczian T, Rodriguez A, Stevens R, Wilke A, Wilkening J, Edwards RA. 2008. The metagenomics RAST server—a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics* 9:386. <http://dx.doi.org/10.1186/1471-2105-9-386>.