

# Draft Genome Sequence of the N<sub>2</sub>-Fixing Cyanobacterium *Nostoc piscinale* CENA21, Isolated from the Brazilian Amazon Floodplain

TIAGO LEÃO,<sup>a,c</sup> PEDRO IVO GUIMARÃES,<sup>b,c</sup> ALINE GRASIELLE COSTA DE MELO,<sup>c</sup> ROMMEL THIAGO JUCÁ RAMOS,<sup>c</sup> PEDRO NUNO LEÃO,<sup>e</sup> ARTUR SILVA,<sup>c</sup> MARLI FATIMA FIORE,<sup>d</sup> MARIA PAULA CRUZ SCHNEIDER<sup>c</sup>

Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California San Diego, La Jolla, California, USA<sup>a</sup>; Department of Biological Systems Engineering, Virginia Tech, Blacksburg, Virginia, USA<sup>b</sup>; Federal University of Pará, Institute of Biological Sciences, Center of Genomic and System Biology, Belém, Pará, Brazil<sup>c</sup>; Center for Nuclear Energy in Agriculture, University of São Paulo, Piracicaba, São Paulo, Brazil<sup>d</sup>; Interdisciplinary Centre of Marine and Environmental Research (CIIMAR/CIIMAR), University of Porto, Porto, Portugal<sup>e</sup>

**We announce here the draft genome sequence of *Nostoc piscinale* CENA21, a diazotrophic heterocyst-forming cyanobacterium isolated from the Solimões River, Amazon Basin, Brazil. It consists of one circular chromosome scaffold with 11 contigs and total size of 7,094,556 bp. Secondary metabolite annotations indicate a good source for the discovery of novel natural products.**

Received 11 February 2016 Accepted 15 February 2016 Published 31 March 2016

**Citation** Leão T, Guimarães PI, de Melo AGC, Ramos RTJ, Leão PN, Silva A, Fiore MF, Schneider MPC. 2016. Draft genome sequence of the N<sub>2</sub>-fixing cyanobacterium *Nostoc piscinale* CENA21, isolated from the Brazilian Amazon floodplain. *Genome Announc* 4(2):e00189-16. doi:10.1128/genomeA.00189-16.

**Copyright** © 2016 Leão 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 Tiago Leão, tferreir@ucsd.edu.

The *Cyanobacteria* are among the most ancient life forms on Earth, dating back approximately 3 billion years ago, and they are responsible for the rise of the atmospheric oxygen on the primitive Earth (1). They are a group of photosynthetic prokaryotes capable of surviving in almost every conceivable environment. The great adaptation skills of cyanobacteria are directly related to their capacity to produce several bioactive natural compounds (2). In particular, the genus *Nostoc* is known for the production of several highly bioactive compounds, such as cryptophycins, nosperin, nostocyclopeptide, nostopeptolide, and nostophycin (3–7), as well as potent toxins, like microcystin and nodularin (8, 9). This genus is also one of the most common dinitrogen (N<sub>2</sub>)-fixing cyanobacteria found in terrestrial, aerial, and aquatic ecosystems (10). However, in contrast to several other cyanobacterial genera, the number of *Nostoc* genomes currently available remains fairly low. Therefore, we announce here the draft genome sequence of *Nostoc piscinale* CENA21, isolated from a floodplain soil of Mari-Mari Island, Solimões River, Amazon Basin, Brazil (02°45'00.00''S, 65°15'00.00''W), in August 1985, as previously described (11), and this strain is maintained in the culture collection of CENA/USP in Piracicaba, São Paulo, Brazil. DNA extraction was performed using a classic phenol-chloroform protocol (12) optimized to cyanobacteria. Genomic DNA was sequenced by both the Ion Torrent PGM platform, using the Ion Torrent 318 Chip, and SOLiD 5500xl, using mate-pair 50-bp reads. Reads were filtered using the FASTX-toolkit, and a minimum threshold of Phred 20 to 50% of the reads' length. The filtered reads from Ion Torrent sequencing were assembled with the MIRA version 4.0 assembler, and the contigs were curated using the Lasergene platform. The assembled contigs were binned by a pipeline adapted from Albertsen et al. (13).

*Nostoc* sp. strain PCC 7107 (with the highest average nucleotide identity to CENA21 of approximately 87.5%) was selected as the reference genome and used to order the binned contigs into scaffolds by the software CONTIGuator, and then the software GapFiller was used to reduce the number of Ns and gaps. The final draft of the *N. piscinale* CENA21 genome consists of a single circular scaffold (11 contigs), total size of 7,094,556 bp, G+C content of 40.54%, and average coverage of 129×. Expert review annotation from the Joint Genome Institute (JGI) IMG/ER database confirms that this strain harbors required genes for photosynthesis, N<sub>2</sub> fixation, heterocyst formation, biosynthesis of most proteinogenic amino acids, and important cofactors. According to the presence/absence of 79 single-copy genes found in 95% of all 44 *Nostocales* genomes currently at the IMG/ER database, we estimate the genome completeness of the announced strain to be approximately 96.2%. The prediction of secondary metabolites using the antiSMASH tool indicates that 8.3% of the assembled genome is dedicated to natural product biosynthesis, grouped into 16 biosynthetic gene clusters (BGCs), including 6 bacteriocins, a nonribosomal peptide-synthetase (NRPS), 3 hybrid NRPS/polyketide synthases (PKS), 3 terpenes, and a lantipeptide. Preliminary analysis indicated that most of those BGCs are cryptic, representing a good source for the discovery of novel natural products.

**Nucleotide sequence accession number.** The *N. piscinale* CENA21 strain draft genome and annotation information were deposited in DDBJ/EMBL/GenBank under the accession no. CP012036.

## ACKNOWLEDGMENTS

This work has been supported by grants from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and from the Funda-

ção de Amparo à Pesquisa do Estado do Pará (BIONORTE-CNPq/FAPESPA) to M.P.C.S.

## FUNDING INFORMATION

This work, including the efforts of Marli Fatima Fiore and Maria Paula Cruz Schneider, was funded by MCTI | Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (307045/2010-2 and 306607/2012-3). This work, including the efforts of Tiago Leão and Pedro Ivo Guimaraes, was funded by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) (13425-13-7 and 13194-13-5). This work, including the efforts of Pedro Nuno Leão, was funded by Ministry of Education and Science | Fundação para a Ciência e a Tecnologia (FCT) (IF/01358/2014).

## REFERENCES

1. Crowe SA, Døssing LN, Beukes NJ, Bau M, Kruger SJ, Frei R, Canfield DE. 2013. Atmospheric oxygenation three billion years ago. *Nature* 501: 535–538. <http://dx.doi.org/10.1038/nature12426>.
2. Dittmann E, Gugger M, Sivonen K, Fewer DP. 2015. Natural product biosynthetic diversity and comparative genomics of the *Cyanobacteria*. *Trends Microbiol* 23:642–652. <http://dx.doi.org/10.1016/j.tim.2015.07.008>.
3. Magarvey NA, Beck ZQ, Golakoti T, Ding Y, Huber U, Hemscheidt TK, Abelson D, Moore RE, Sherman DH. 2006. Biosynthetic characterization and chemoenzymatic assembly of the cryptophycins. Potent anticancer agents from cyanobionts. *ACS Chem Biol* 1:766–779. <http://dx.doi.org/10.1021/cb6004307>.
4. Kampa A, Gagunashvili AN, Gulder TAM, Morinaka BI, Daolio C, Godejohann M, Miao VPW, Piel J, Andrésson ÖS. 2013. Metagenomic natural product discovery in lichen provides evidence for a family of biosynthetic pathways in diverse symbioses. *Proc Natl Acad Sci USA* 110: E3129–E3137.
5. Becker JE, Moore RE, Moore BS. 2004. Cloning, sequencing, and biochemical characterization of the nostocyclopeptide biosynthetic gene cluster: molecular basis for imine macrocyclization. *Gene* 325:35–42. <http://dx.doi.org/10.1016/j.gene.2003.09.034>.
6. Hoffmann D, Hevel JM, Moore RE, Moore BS. 2003. Sequence analysis and biochemical characterization of the nostopeptolide A biosynthetic gene cluster from *Nostoc* sp. GSV224. *Gene* 311:171–180. [http://dx.doi.org/10.1016/S0378-1119\(03\)00587-0](http://dx.doi.org/10.1016/S0378-1119(03)00587-0).
7. Fujii K, Sivonen K, Kashiwagi T, Hirayama K, Harada KI. 1999. Nostophycin, a novel cyclic peptide from the toxic cyanobacterium *Nostoc* sp. *J Org Chem* 152:.
8. Gehringer MM, Adler L, Roberts AA, Moffitt MC, Mihali TK, Mills TJ, Fieker C, Neilan BA. 2012. Nodularin, a cyanobacterial toxin, is synthesized in planta by symbiotic *Nostoc* sp. *ISME J* 6:1834–1847.
9. Fewer DP, Wahlsten M, Österholm J, Jokela J, Rouhiainen L, Kaasalainen U, Rikkinen J, Sivonen K. 2013. The genetic basis for O-acetylation of the microcystin toxin in cyanobacteria. *Chem Biol* 20: 861–869. <http://dx.doi.org/10.1016/j.chembiol.2013.04.020>.
10. Komárek J 2013. 19/3: Cyanoprokaryota; Teil 3: heterocystous genera. Springer, Berlin, Germany.
11. Fiore Mde F, de F, Neilanc BA, Copp JN, Rodrigues JLM, Tsaia SM, Lee H, Trevors JT. 2005. Characterization of nitrogen-fixing cyanobacteria in the Brazilian Amazon floodplain. *Water Res* 39:5017–5026. <http://dx.doi.org/10.1016/j.watres.2005.10.002>.
12. Sambrook J, Russell DW. 2001. *Molecular cloning: a laboratory manual*, vol 1. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
13. Albertsen M, Hugenholtz P, Skarshewski A, Nielsen KL, Tyson GW, Nielsen PH. 2013. Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. *Nat Biotechnol* 31:533–538.