# Draft Genome Sequence of *Blastomonas* sp. Strain CACIA 14H2, a Heterotrophic Bacterium Associated with Cyanobacteria

**Alex Ranieri Jerônimo Lima,a,b Andrei Santos Siqueira,a Bruno Garcia Simões dos Santos,a Fábio Daniel Florêncio da Silva,a Clayton Pereira Lima,b Jedson Ferreira Cardoso,b João Lídio S. G. Vianez-Júnior,b Márcio Roberto Teixeira Nunes,b Evonnildo Costa Gonçalvesa,b**

Laboratório de Tecnologia Biomolecular, Instituto de Ciências Biológicas (ICB), Universidade Federal do Pará (UFPA), Belém, Pará, Brazila; Centro de Inovações Tecnológicas (CIT), Instituto Evandro Chagas (IEC), Ananindeua, Pará, Brazilb

**With the new methods for assembling sequence data from metagenomic samples, the genomic study of heterotrophic bacterium-cyanobacterium associations can now be improved. In this work, the draft genome sequence of *Blastomonas* sp. strain CACIA 14H2, obtained from a nonaxenic culture of a *Cyanobium* sp., is presented.**

Address correspondence to Alex Ranieri Jerônimo Lima, alexranieri@hotmail.com.

The development of new methods of assembling sequencing data from metagenomic samples (1, 2) means it has become possible to address via genome sequences the unique features shared by heterotrophic bacterium-cyanobacterium associations (3) that are capable of producing several compounds of biotechnological interest (4, 5). Thus, we obtained the draft genome of *Blastomonas* sp. strain CACIA 14H2 from the sequencing of a nonaxenic culture of cyanobacteria. The genus *Blastomonas* is characterized by rod-shaped Gram-negative bacteria that produce bacteriochlorophyll $\alpha$ under aerobic conditions (6, 7). So far, public databases have presented only one draft genome sequence for the genus (8).

The cyanobacterial isolate was obtained from a water sample collected in December 2010 from the Tucuruí hydroelectric dam (3°49′55″S, 49°38′50″W), Pará, Brazil. DNA samples were obtained from *Cyanobium* sp. strain CACIA 14 cultured 6 months apart. The two nonpaired libraries (9) were sequenced using the GS FLX 454 sequencer (Roche Life Science), resulting in 660,228 (~255 Gbp) and 815,325 (~357 Gbp) reads for the first and second runs, respectively. Both datasets were assembled separately with Newbler 2.6 (minimum read size, 45 bp; minimum overlap, 40 bp; minimum overlap identity, 90%). These analyses generated 3,654 and 3,256 contigs >1 kbp in length, with $N_{50}$ values of 2,149 bp and 37,998 bp. The assembled contigs were identified and separated using a metagenomic assembling pipeline (1) for each putative organism. The assembled contigs from the second run were used to determine the genome coverage. The draft genome of *Blastomonas* sp. CACIA 14H2 contains 72 contigs (4,067,409 bp), ranging from 5,698 to 260,673 bp; the average coverage was estimated to be 25×, the $N_{50}$ is 89,796, and the G+C content is 65.24%. The pipeline uses 107 hidden Markov models for essential genes present in a single copy in 95% of all bacteria. The draft contains 106 of these genes, with PF00750 in duplicate, which normally occurs this way. The structural annotation was performed with the PGAP tool, available on the NCBI website (10),

resulting in 3,787 annotated coding sequences (CDSs) and 42 tRNA genes. The rRNA clusters were predicted by the RNAmmer tool (11). The 16S rRNA predicted is 99% similar to the relative sequence of *Blastomonas natatoria* strain DSM 3183, while the *rpoB* gene is 97% similar to the relative sequence of *Blastomonas* sp. strain AAP53.

The information contained in the draft genome of *Blastomonas* sp. CACIA 14H2 presented here may be useful for elucidating attributes important to heterotrophic bacterium-cyanobacterium associations.

**Nucleotide sequence accession numbers.** This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. AYSC00000000. The version described in this paper is version AYSC01000000.

## REFERENCES

1. **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. http://dx.doi.org/10.1038/nbt.2579.
2. **Nederbragt AJ, Rounge TB, Kausrud KL, Jakobsen KS.** 2010. Identification and quantification of genomic repeats and sample contamination in assemblies of 454 pyrosequencing reads. Sequencing **2010:**782465.
3. **Lupton FS, Marshall KC.** 1981. Specific adhesion of bacteria to heterocysts of *Anabaena* spp. and its ecological significance. Appl. Environ. Microbiol. **42:**1085–1092.
4. **Abed RM, Dobretsov S, Sudesh K.** 2009. Applications of cyanobacteria in biotechnology. J. Appl. Microbiol. **106:**1–12. http://dx.doi.org/10.1111/j.1365-2672.2008.03918.x.
5. **Rastogi RP, Sinha RP.** 2009. Biotechnological and industrial significance of cyanobacterial secondary metabolites. Biotechnol. Adv. **27:**521–539. http://dx.doi.org/10.1016/j.biotechadv.2009.04.009.

6. **Sly LI, Cahill MM.** 1997. Transfer of *Blastobacter natatorius* (Sly 1985) to the genus *Blastomonas* gen. nov. as *Blastomonas natatoria* comb. nov. Int. J. Syst. Bacteriol. **47:**566–568. http://dx.doi.org/10.1099/00207713-47-2-566.

7. **Hiraishi A, Kuraishi H, Kawahara K, .** 2000. Emendation of the description of *Blastomonas. natatoria* (Sly 1985) Sly and Cahill 1997 as an aerobic photosynthetic bacterium and reclassification of *Erythromonas ursincola* Yurkov et al. 1997 as *Blastomonas ursincola* comb. nov. Int. J. Syst. Evol. Microbiol. **50**(Pt **3**):1113–1118.

8. **Zeng Y, Feng F, Liu Y, Wu Z, Jian J.** 2013. Whole-genome sequences of an aerobic anoxygenic phototroph, *Blastomonas* sp. strain AAP53, isolated from a freshwater desert lake. Genome Announc. **1**(2):e00071-13. http://dx.doi.org/10.1128/genomeA.00071-13.

9. **Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Ho CH, Irzyk GP, Jando SC, Alenquer ML, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt Ka, Volkmer Ga, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, Rothberg JM.** 2005. Genome sequencing in microfabricated high-density picolitre reactors. Nature **437:**376–380.

10. **Angiuoli SV, Gussman A, Klimke W, Cochrane G, Field D, Garrity G, Kodira CD, Kyrpides N, Madupu R, Markowitz V, Tatusova T, Thomson N, White O.** 2008. Toward an online repository of Standard Operating Procedures (SOPs) for (meta)genomic annotation. Omics **12:**137–141. http://dx.doi.org/10.1089/omi.2008.0017.

11. **Lagesen K, Hallin P, Rødland EA, Staerfeldt HH, Rognes T, Ussery DW.** 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. **35:**3100–3108. http://dx.doi.org/10.1093/nar/gkm160.