

TITLE: CHLOROPLAST GENOME OF AN AMAZONIAN ISOLATED ACIDOPHILIC *HEVEOCHLORELLA* SP.

AUTHORS: SIMÕES, B.G.¹; PIRES, E.S.¹; GUIMARÃES, J.T.¹; OLIVEIRA, G.¹; DALL'AGNOL, H.M.B.².

INSTITUTIONS: 1) INSTITUTO TECNOLÓGICO VALE, BELÉM, PA (R. BOAVENTURA DA SILVA, 955, CEP 66055-090, BELÉM – PA, BRAZIL); 2) UNIVERSIDADE FEDERAL DO MARANHÃO, SÃO LUIS, MA (AV. DOS PORTUGUESES, 1966, CEP 65065-545, SÃO LUÍS - MA, BRAZIL).

ABSTRACT:

Chloroplast genomes (cpDNA) are a valuable source of information about the phylogenetic relationships among green algae, due to its high content and condensed distribution of genes when compared to nuclear genome. Furthermore, they have been providing breakthroughs to the industry with insights about the organisms' biotechnological potential. However, the Amazonian algal biodiversity is still poorly studied, mostly when compared to higher plants. In this sense, this study present the first chloroplast genome of an Amazonian isolated acidophilic *Heveochlorella* sp., a little-known genus of the Chlorellaceae family. It is a non-motile and spherical unicellular freshwater green microalga, from a natural lake located in a region with iron ore mining activity, in the southern part of the Serra dos Carajás, Pará State, Brazil. Cultivation was performed in BG-11 media (pH 2.5), total DNA was extracted using MOBIO UltraClean® microbial DNA isolation kit and sequenced in a paired-end Illumina MiniSeq library and in an Ion Torrent Library. The reads were quality filtered using the FASTX-Toolkit collection and assembled using two different softwares (NOVOPlasty and SPAdes). The assemblies were merged using Geneious R9. The automatic annotation were made with CpGAVAS and manually curated using Artemis and NCBI Blast. The final annotated genome were compared to the other genomes, of the same family, available at NCBI database. The *Heveochlorella* sp. cpDNA sequence assembled as a circular map of 131,273 base pairs (bp), with an GC content of 34.4%, similarly to the others Chlorellaceae strains. It encodes 70 genes, including 42 protein coding genes, two rRNAs and 26 tRNAs. The Chlorellaceae family chloroplast genomes encode a higher number of genes (111 to 210), while genome size is not uniform among them (84k to 236k bp). Like others strains into the family, *Heveochlorella* sp. genome does not show an architecture composed of a large single-copy region (LSC) and a small single-copy separated by a pair of inverted repeats (IRs). The sequencing of the first chloroplastial genome from an Amazonian isolated acidophilic *Heveochlorella* sp. is the first step into the understanding of the local microalgal biodiversity evolutionary history, their adaptations to changing environment etc. Those basic knowledge can aid the forthcoming works to deeply explore its striking features, as grow at a very low pH (< 3), at the development of sustainable biotechnological applications.

Keywords: Chlorophyta; Brazilian Amazon; mining activity; chloroplast genome; acidophilic microalgae.

Development Agency: Vale Company, CNPq (459913/2014-04), Bruno Garcia Simões was funded by CNPq (443270/2015-5).