

**TITLE:** PHYLOGENETIC RESOLUTION OF GENES ENCODING 16S rRNA, 23S rRNA AND AAA-METALLOPROTEASES FOR CIANOBACTERIAL SPECIES DELINEATION

**AUTHORS:** CRUZ, F. J. G<sup>1</sup>; SANTANA, L. L<sup>1</sup>; CRUZ, D. C. B<sup>1</sup>; FIGUEIREDO, C. N<sup>1</sup>; SOUZA, J. T<sup>2</sup>; MARBACH, P. A. S<sup>1</sup>

**INSTITUTIONS:** 1 - UNIVERSIDADE FEDERAL DO RECÔNCAVO DA BAHIA, BAHIA, BA (R. RUI BARBOSA, 710, CRUZ DAS ALMAS, CEP 44380-000, BAHIA – BA, BRAZIL) 2 - UNIVERSIDADE FEDERAL DE LAVRAS, MINAS GERAIS, MG (AV. DOUTOR SYLVIO MENICUCCI, 1001, LAVRAS - MG, CEP 37200-000)

**ABSTRACT:**

Despite its environmental and biotechnological importance, the taxonomy and delineation of cyanobacterial species remains a difficult task. Usually, genes encoding 16S and 23S rRNA are used to infer evolutionary relationships of species in cyanobacteria. However, the existence of multiple copies of these genes in cyanobacterial genomes contributes to the introduction of bias in phylogenetic inferences when they are not inherited vertically during speciation. The single-copy gene AAA-metalloprotease is ubiquitous in bacterial genomes and may have potential in cyanobacterial taxonomy. The objective of this work was to compare the genes encoding 16S, 23S rRNA and AAA-metalloprotease in the delimitation of cyanobacterial species. The BLASTP and TBLASTN tools were used to recover 16S, 23S rRNA and AAA-metalloprotease from the cyanobacterial genomes available at the NCBI database. A total of 222 sequences of 16S rRNA, 223 of 23S rRNA and 411 from AAA-metalloproteases were recovered. These sequences were used to perform phylogenetic reconstructions in the MEGA 6.06 software using the Maximum Likelihood method. The number of gene copies encoding 16S and 23S rRNA per genomes ranged from 1 copy, in *Calothrix* sp. 336/3 to 5 copies in *Anabaena* sp. 90, *Anabaena* sp. Wa102 and *Nostoc* sp. PCC 7524. The topology of the 16S rRNA Maximum likelihood tree indicates that the copies encoded in the same genome are phylogenetically more related to each other indicating that they originated in events of gene duplications. Similar results were observed in the phylogeny of genes encoding 23S rRNA. Nine clades of the 16S rRNA Maximum likelihood tree were polytomic and they contained 76 genes. In contrast, only two polytomic clades were found in the 23S rRNA maximum likelihood tree and these clades contained 18 genes. The molecular phylogeny of cyanobacterial AAA-metalloproteases indicates that there is only one isoform present in every studied genome, suggesting that they are orthologs. The topology of the clade containing these genes, here called cb3 group, has no polytomy. The results show that the genes encoding cyanobacterial AAA-metalloproteases from the cb3 group have higher phylogenetic resolution than the genes encoding cyanobacterial 16S and 23S rRNA. Therefore, we conclude that the cyanobacterial AAA-metalloproteases cb3 can be used as markers for the delineation of cyanobacterial species.

**Keywords:** FtsH, bacterial taxonomy, phylogeny

**Development agency:** Universidade Federal do Recôncavo da Bahia