

# Towards a simpler photoautotrophic cell: conserved and variable genes in *Synechococcus elongatus*

Luis Delaye<sup>1</sup>, Carmen M. González-Domenech<sup>2,3</sup> and Andrés Moya<sup>3</sup>

<sup>1</sup> CINVESTAV-Irapuato, México

<sup>2</sup> Facultad de Farmacia, Universidad de Granada

<sup>3</sup> Instituto Cavanilles de Biodiversidad y Biología Evolutiva, Universitat de València  
Andres.Moya@uv.es

## Abstract

Simpler biological systems should be easier to understand and engineer. One way to achieve biological simplicity is through genome minimization. Here we have looked for genomic islands in the fresh water cyanobacterium *Synechococcus elongatus* PCC 7942 that could be used as targets for deletion for genome minimization. By using a combination of methods we have identified 184 genes that have been horizontally transferred into the genome of *S. elongatus* plus 127 ORFans (Figure 1). These genes have a combination of: a) unusual G+C content; b) unusual phylogenetic similarity; and/or c) a small number of a highly iterated palindrome 1 (HIP1) motif plus an unusual codon usage. We have also corroborated the existence of the largest genomic island by its lack of coverage among metagenomic sequences from a fresh water microbialite. Interestingly, most genes coding for proteins with a diguanylate cyclase domain are predicted to be xenologous, suggesting a role for horizontal gene transfer in the evolution of sensory systems in this cyanobacteria. In parallel we have identified 1401 highly conserved genes that might be essential for cell survival and should not be deleted. These two datasets (variable and conserved genes) comprises ~11.8% and 53.6% of annotated genes in *S. elongatus*. Our results set a guide to non-essential genes in *S. elongatus* PCC 7942 indicating a path towards the engineering of a simpler photoautotrophic cell.

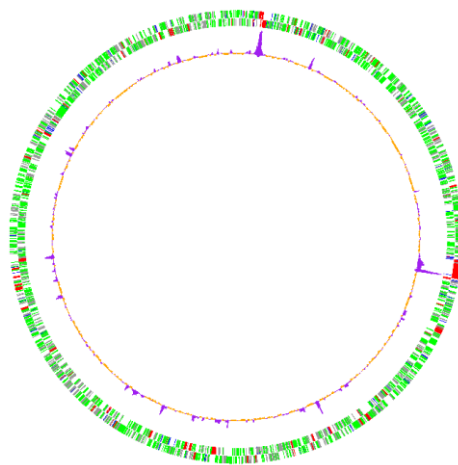


Figure 1. Conserved and variable regions in the genome of *S. elongatus* PCC 7942. Outer circle. Red: variable genes; green: conserved genes; gray: other. Inner circle. Regions of atypical tri-nucleotide composition are shown in purple.