120. Analysis of the Transcriptional Response of Synechococcus sp. PCC 7002 to Specific Growth Conditions from a Compendium of RNA-seq Data

Ryan S. McClure[1], Lye Meng Markille[1], Ronald C. Taylor[1], Jason E. McDermott[1], Marcus Ludwig[2], Donald A. Bryant[2], **Allan Konopka[1]**, and **Alex S. Beliaev[1]**

[1]Biological Sciences Division, Pacific Northwest National Laboratory, Richland, WA 99352
[2]Department of Biochemistry and Molecular Biology, The Pennsylvania State University, State College, PA 16802

Project Goals: The regulatory machinery of cyanobacteria, which evolved to provide ecophysiological advantages, plays a major role in both short- and long-term adaptations to environmental perturbations. The challenges of extreme environments require coordinated adjustment of photosynthetic efficiency, carbon processing rates, and "downstream" carbon partitioning between biomass and metabolite storage pools. Understanding how regulatory controls dynamically integrate the external inputs to produce intracellular adaptations will inform strategies to overcome productivity constraints and optimize metabolism for biofuels production. Availability of a well-established genetic system for *Synechococcus* sp. PCC 7002, in conjunction with high-throughput 'omics approaches, provides the scale and resolution for a comprehensive analysis of transcriptional regulation in cyanobacteria.

Synechococcus sp. PCC 7002 (hereafter *Synechococcus* 7002) is a fast-growing unicellular cyanobacterium that can be found in the brackish or saline water of tidal estuaries, a dynamic environment where strict gene regulatory mechanisms are necessary for survival. Such regulation takes place with fewer transcription factors than in proteobacterial systems, and in the presence of numerous antisense transcripts, suggesting post-transcriptional methods of regulation. High resolution quantitative analyses of RNA *via* next-generation sequencing offers the ability to analyze changes, both in specific gene products and the organism's transcriptome as a whole, when exposed to different stimuli. Such analyses represent an alternative to genome analysis alone when identifying regulons as co-expressed genes can be identified regardless of genomic context. Deep sequencing technologies also provide the ability to analyze changes in both known open reading frames (ORFs) as well as unannotated RNA transcripts such as regulatory small RNAs (sRNAs). To that end, widespread occurrence of non-coding antisense transcripts (asRNA) in the genomes of cyanobacteria suggests a prevalence of post-transcriptional regulation. As previously postulated, the regulation of antisense transcription is likely to be tailored to its mode of action, while the co-expression patterns between asRNAs and their targets might indicate the mechanism of action.

In this study, we combined RNA-seq analysis from 41 different growth conditions to determine how *Synechococcus* 7002 responds on a global level using cluster analysis and context likelihood of relatedness (CLR) approaches. Clusters of co-expressed genes were generated and, in conjunction with functional enrichment analysis, were used to generate an overview of regulatory and metabolomic pathways in *Synechococcus* 7002. As RNA-seq also allows for the determination of 5' untranslated regions (UTRs) these were compared within clusters of co- expressed genes to identify possibly homology indicating common transcription factors or sRNAs regulating such gene clusters. In addition, several hundred UTRs were examined to gain insight into the specific transcriptional regulatory mechanisms of *Synechococcus* 7002. Several UTRs were found to be altered based on growth conditions. Comparing N-limitation to C-limited conditions showed UTR changes of at least 30 nucleotides in 45 genes and similar changes were found when comparing other growth conditions. Our analysis was not

limited to protein coding genes as we also identified over 450 instances of unannotated transcription either within an intergenic region or opposite a known protein-coding gene. Through a stepwise application of known sRNA characteristics we describe 24 sRNAs that are between 30-250 nucleotides in length and contain a predicted Rho Independent Terminator (RIT). Expression of several of these sRNAs is regulated based on specific environmental conditions and through an analysis of homologous base-pairing and concordant expression patterns determined from RNA-seq data we propose several mRNA targets of these sRNAs. For example, a sRNA at 1152089-1152190 shows a 2 fold increase in expression when cells are grown at 0.D. 0.4 compared to O.D. 0.1.

The sequence of this sRNA displays homologous basepairing to SYNPCC7002_A0916, a hypothetical protein that shows a 2 fold *decrease* in expression under these conditions. Such putative interactions between RNA strands suggests that sRNA may be a negative regulator of SYNPCC7002_A0916. These studies are the first to combine a large number of growth conditions to gain a global perspective on the transcriptome of *Synechococcus* 7002 as it responds to varying environmental stimuli. Knowledge gained regarding regulatory patterns of this marine organism will be invaluable both from an environmental perspective as well as in the context of bioenergy and biotechnology applications using cyanobacteria.