

Computational Inference of Transcription Regulation and Response Networks in *Synechococcus* sp. WH8102

Zhengchang Su, PhuongAn Dam, Victor Olman, Ying Xu
Department of Biochemistry and Molecular Biology
University of Georgia at Athens, USA, and
Institute of Computational Biology
Oak Ridge National Laboratory
Oak Ridge, TN 37831, USA

Abstract

We have recently developed a computational capability for (a) inference of individual response networks and (b) global transcription regulation networks in *Synechococcus* sp. WH8102. The overall framework for inference of response networks can be summarized as follows. The framework employs a template-based network inference strategy. Under this framework, we first identify organisms that are known to have similar response systems to *Synechococcus* sp for a particular stimulus, and that have a large quantity of experimental data available, particularly microarray gene expression data. For each such “template” organism, we first build network models, and then map these template network models to *Synechococcus* sp. Through literature search and database search, we build a conceptual framework first for a specific response network in each template organism. The detailed network models will be then built through a series of predictions of gene functions, operon and regulon structures, transcription regulatory binding sites, protein-protein interactions and protein-DNA interactions, and incorporate these prediction results into the conceptual model. After the template network models are built, we map these models to *Synechococcus* sp through a constrained orthologous gene mapping scheme. The basic idea is to map genes of a network model to their predicted homologous genes under the constraints that these genes should belong to a group of co-regulated operons, which is predicted based on operon structure prediction and prediction of transcriptional regulatory binding sites. The mapped network models will then be refined and expanded through application of a genome-scale protein-protein interaction network, which is predicted through a series of computational prediction tools. We have applied this capability to derive a number of response networks in *Synechococcus* sp, such as phosphorus assimilation pathway, carbon fixation pathway, nitrogen assimilation pathway. Experiments are currently under way to validate these predictions. A major challenge remains in automated mining of literature to extract useful information to build the initial conceptual network to collect information about gene functions and interactions.

Prediction of global transcription regulatory network is done through prediction of (a) transcription factors (TF) at genome scale, (b) operon predictions, (c) prediction of TF binding sites for each predicted operon, and (d) prediction of interactions between predicted TFs and TF binding sites. The key techniques we have used for many of these predictions include (i) comparative genome analysis and (ii) computational inference of co-evolution relationships through applications of methods such as phylogenetic profile analysis. Through these predictions, we have identified a number of global transcription regulators and a sequence of transcription regulators responsive to particular stimuli. Interaction relationship between these TFs and the predicted TF binding sites is predicted based on predicted co-evolution relationships of TFs and genes in each predicted operon. A preliminary global network of transcription regulation is predicted.

Acknowledgement

This work is support by the US Department of Energy’s Genomes to Life program (www.deogenomestolife.org) under project, “Carbon Sequestration in *Synechococcus* sp.: From Molecular Machines to Hierarchical Modeling” (www.genomes-to-life.org). YX’s work is also supported, in part, by National Science Foundation (# NSF/DBI-0354771).