Identification of Transcription Factor Binding Sites and Characterization of Promoter Architecture in the model Diatom *Phaeodactylum tricornutum*

Andrew E. Allen[1,2*], Sarah R. Smith, Hong Zheng, Pratap Venepally

[1]Microbial and Environmental Genomics, J. Craig Venter Institute, La Jolla, CA 92037, USA; [2]Scripps Institution of Oceanography, Integrative Oceanography Division, University of California, San Diego, La Jolla, CA 92093, USA.

Project Goals:

We propose integration of genome-scale modeling with genome engineering to optimize energy and metabolte flux through subcellular compartments to promote efficient production of high value and fuel-related metabolites. Through the proposed research activities, we aim to construct streamlined artificial chromosomes encoding reprogrammed biological modules designed for *in vivo* optimization of electron flow efficiency, photosynthesis, and overall cellular growth while directing key metabolic precursors away from storage carbohydrates and into lipids or branched chain amino acids (BCAA). The underlying goal of the proposed research is to produce strains of diatoms encoding cellularly compartmentalized biosynthesis pathways on an artificial chromosome, with the natural genetic background altered to include knockouts of respective native genes as well as the installation of *in vivo* metabolite bioreporters.

Abstract:

Transcription factors (TFs) regulate gene expression by binding DNA in gene promoters and have an important role in activating the cellular response to shifting environmental conditions. To date, several diatom transcriptome studies have shown that suites of genes are co-expressed in response to shifting conditions (i.e. nutrients or light), identifying putative gene regulons. However, little is known about the roles or binding sites of specific TFs that elicit these transcriptional responses as only a few TFs have been characterized in diatoms. We investigated transcription factor binding sites, using a combination of bioinformatics-based promoter analysis and high-throughput in vitro DNA affinity purification sequencing (DAP-seq). DAP-seq enables genome-wide characterization of transcription factor binding sites. We identify a nitrate-response element enriched in the promoters of nitrate assimilation genes that is similar to the binding site for the human transcription factor ETS. We also characterize the optimal binding sites for different genes in the bZIP transcription factor family from *P. tricornutum*, facilitating description of the genes they regulate. Specific knowledge of promoter architecture is valuable for the development of tools for molecular investigations and genetic engineering of diatoms and is essential in order to understand how reprogramming of gene expression is accomplished to achieve appropriate cellular responses to environmental signals.

Supported by DE-SC0008593*, Biological and Environmental Research*