Title: Mapping transcription factor-mediated remodeling of diatom metabolism in response to shifting environmental conditions

Authors: Sarah R. Smith[1], Hong Zheng[1], Pratap Venepally[1], and **Andrew E. Allen[1,2]* (aallen@jcvi.org)**

Institutions: [1]J. Craig Venter Institute, La Jolla, California; [1]Scripps Institution of Oceanography, University of California, San Diego

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. Specific goals and technical approaches are focused around four themes: 1) Modeling and Flux studies, 2) Photosynthetic efficiency, 3) Linking metabolic and regulatory networks, 4) Genome scale engineering methodology and application.

Abstract text: Transcription factors regulate gene expression by binding DNA and promoting (activate) or block (repress) recruitment of RNA polymerase. Therefore, they are a key part of regulating the enaction of gene expression programs that govern regular growth (progression through the cell cycle) and allow cells to respond and acclimate to shifts in environmental conditions (such as light, and nutrient status) to optimally maintain homeostasis and survive. Despite the importance of diatoms in the marine environment, and their relevance for biofuels development, little is known about these mechanisms with any molecular detail. DNA affinity purification sequencing (DAP-Seq) is a high-throughput *in vitro* method to characterize transcription factor binding sites genome-wide. Using this method, we have mapped the transcription factor binding sites for >100 transcription factors in the *P. tricornutum* genome. Our results corroborate findings from previous investigators for the few transcription factors that have been functional characterized, and greatly expand the catalog of these key molecular components of signal transduction cascades in diatoms. Notably, we identify a homolog of the fungal regulator of the nitrate regulon and link it to the regulation of nitrate metabolism in *P. tricornutum* and across diatom diversity. Further, we identify transcription factors involved in other key cellular events such as the coordination of metabolic shifts across diel cycles, and activation of the carbon concentrating mechanism. In diatoms, heat shock factors (HSF) have expanded massively into diatom-specific distinct classes, and we detail for the first time, variation in the binding sites of the different HSF classes providing insight into the functional

significance of their evolution in this group. Through linking the prevalence of binding sites in the promoters of condition-specific regulons, we develop the first genome-scale view of the metabolic regulatory network in diatoms.

Funding statement: This research was supported by the DOE Office of Science, Office of Biological and Environmental Research (BER), grant no. DE-SC0018344